PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/57, 9/52, 15/76, 1/21, 1/15, C07K 5/06, C12N 15/62, C07K 14/47, C12P 21/06, G01N 33/535 // C12N 1/21, C12R 1:465, C07K 14/54, 14/475, 14/535, 14 /505

A2

(11) International Publication Number: **WO 95/17512**

(43) International Publication Date: 29 June 1995 (29.06.95)

(21) International Application Number: PCT/US94/14772

(22) International Filing Date: 22 December 1994 (22.12.94)

(30) Priority Data:
08/173,508 23 December 1993 (23.12.93) US

(60) Parent Application or Grant
(63) Related by Continuation
US 08/173,508 (CIP)
Filed on 23 December 1993 (23.12.93)

(71) Applicant *(for all designated States except US)*: CANGENE CORPORATION [CA/CA]; 6280 Northwest Drive, Mississauga, Ontario L4V 1J7 (CA).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: BARTFELD, Daniel [IL/CA]; 89 Overlook Place, North York, Ontario M3H 4P5 (CA). BUTLER, Michael, J. [GB/GB]; Worts Causeway, Cambridge CB1 4RN (GB). HADARY, Dany [IL/CA]; 40 Maryvale Crescent, Richmond Hill, Ontario L4C 6P8 (CA). JENISH, David, L. [CA/CA]; 4104 Fieldgate Drive, Mississauga, Ontario L4W 2C4 (CA). KRIEGER, Timothy, J. [US/CA]; 24 Jameson Crescent, Brampton, Ontario L6S 3W3 (CA). MALEK, Lawrence, T. [US/CA]; 3 Viewmount Crescent, Brampton, Ontario L6Z 4P4 (CA). WALCYZK, Eva [CA/CA]; 6037 Childham Crescent, Mississauga, Ontario L5N 2R8 (CA). SOOSTMEYER, Gisela [CA/CA]; 166 Hedgerow Lane, Kleinburg, Ontario L0G 1C0 (CA).

(74) Agents: BENT, Stephen, A. et al.; Foley & Lardner, Suite 500, 3000 K Street, N.W., Washington, DC 20007-5109 (US).

(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LT, LU, LV, MD, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: PROTEASES FROM STREPTOMYCES AND USE THEREOF IN PROTEIN EXPRESSION SYSTEMS

(57) Abstract

A family of proteases endogenous to *Streptomyces* cells degrades exogenous proteins secreted from *Streptomyces* host cells. The previously unidentified proteases include (1) tripeptidyl aminopeptidase designated "Tap", (2) tripeptidyl aminopeptidase designated "Ssp", (3) X-Pro-Metalloendoproteinase designated "XP-Mep", and (4) other proteases derived from *Streptomyces* which degrade certain substrates under certain conditions. Degradation was alleviated by selective inhibition of secreted proteases or by using improved strains which lack or have impaired degradation proteases. An irreversible inhibitor was designed based upon the mechanism and substrate specificity of the target protease. Hosts expressing proteases were also produced. Uses of the proteases include immunoassays and proteolytic removal of peptides and polypeptides to improve secretion of exogenous proteins.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

PROTEASES FROM STREPTOMYCES AND USE THEREOF IN PROTEIN EXPRESSION SYSTEMS

BACKGROUND OF THE INVENTION

This invention relates to improved *Streptomyces* strains which lack degradative proteases. The strains are used for expression of peptides and polypeptides. The invention also relates to inhibitors of such proteases and uses of such proteases. New *Streptomyces* proteases and their genetic coding sequences are presented. Uses of the proteases and impaired proteases are described.

Genetic expression systems generally consist of host cells encompassing a gene to be expressed, and vectors which introduce the gene into the host cells. Under conditions allowing expression, the host cells make a product, generally a protein.

A problem is that host cells have a variety of endogenous proteases which degrade proteins expressed by the systems. These proteases, of which aminopeptidases, dipeptidyl aminopeptidases, tripeptidyl aminopeptidases and endoproteases are examples, interfere with the commercial use of genetic expression systems by degrading proteins expressed by the systems in different ways.

Aminopeptidases remove N-terminal amino acids from proteins. Examples of aminopeptidases include leucine aminopeptidase (Hanson and Frohne, 1976) and aminopeptidase P (Yoshimoto, *et al.*, 1988).

Dipeptidyl aminopeptidases remove N-terminal dipeptides including X-Pro dipeptides from proteins (Lloyd, *et al.*, 1991; Fukusawa and Harada, 1981). Tripeptidyl aminopeptidases remove N-terminal tripeptides from proteins. Tripeptidyl aminopeptidases have not previously been reported in prokaryotes; however, they have been reported in eukaryotes (McDonald, *et al.*, 1985; Balow, *et al.*, 1986; Tomkinson and Jonsson, 1991).

Endoproteases degrade secreted proteins by cleaving amino acid bonds within proteins. Examples of

endoproteases include serine proteases and metalloproteases. They exist in various microbial species and have a wide variety of cleavage site specificities, including specificities adjacent to positively charged, negatively charged, and aromatic amino acids.

Proteases may be neutralized by various methods. One method is to use inhibitors to prevent the degradation of proteins during their purification. This method has been used for proteins derived from yeast and higher eukaryotes and from inclusion bodies derived from *E. coli*. Inhibitors employed in this manner include leupeptin, EDTA, phenylmethanesulfonylfluoride and pepstatin.

But protease inhibitors can harm a living organism. EDTA increases the fragility of many microorganisms and can cause cell lysis. Some inhibitors may be taken up by the organism possibly causing cell death or disrupting cellular functions. Ideally, a protease inhibitor should (1) be soluble in the fermentation media, (2) inhibit the target protease as selectively as possible, (3) not inhibit cell growth, and (4) be cost-effective.

Chloromethylketones are known to provide selective inhibition of some proteases. The earliest studied chloromethylketones, tosyl-L-lysine chloromethylketone (TLCK) and tosylphenylalanine chloromethylketone (TPCK), selectively inhibit trypsin and chymotrypsin, respectively (Schoellman, *et al.*, 1963; Shaw, *et al.*, 1965).

Another method to impair proteases is to use improved strains with impaired proteases to prevent degradation of proteins during production. Improved strains carrying deletional mutations in multiple protease-encoding genes have been made in *Bacillus* strains (Sloma, *et al.*, 1992; PCT/US92/01598 of Omnigene, Inc.).

International Application Number PCT/US92/05532 of Amgen Inc. entitled "Isolation and Characterization of a Novel Protease from *Streptomyces lividans*" describes a protease called "Protease X" of *S.lividans*, and a strain of *S.lividans* deficient in such protease. However, the strain deficient in Protease X does not significantly improve the commercial production of exogenous proteins secreted from *Streptomyces*.

Endogenous proteases have deleteriously affected the quality, quantity or stability of proteins expressed in a *Streptomyces* recombinant genetic expression system designated CANGENUS™. This expression system has been used to ferment and produce a variety of proteins including therapeutic proteins such as granulocyte macrophage-colony stimulating factor (GM-CSF), interleukin-3 (IL-3), interleukin-6 (IL-6), and erythropoietin (EPO) (see Canadian Patent No. 1,295,563, No. 1,295,566 and No. 1,295,567, and U.S. Patent No. 5,200,327) and may be used to produce other proteins including tumor necrosis factor (TNF), stem cell factor (SCF), interleukin -7 (IL-7) and interleukin-2 (IL-2).

To improve commercial production of proteins, a need exists to impair the action of endogenous *Streptomyces* proteases in *Streptomyces* expression systems. Furthermore, a need exists for additional commercial uses of these proteases. It would be helpful if these proteases could be used: (1) in processing recombinant proteins by removing propeptides or fused polypeptides incoporporated into a genetic construct to faciliate expression and secretion to prevent degradation by other proteases, to faciliate purification of polypeptides, and other suitable products; (2) in removing blocked N-terminals from proteins; (3) in digesting connective tissue proteins; (4) in an immunoassay to enable the removal of one or more amino acids from a protein; (5) in a coupled assay for aminopeptidases to remove peptide

blocking groups. The specificity of proteases of the present incitation as described in this patent application enables those skilled in the art to carry out such uses.

Exogenous protein secretion in bacterial cells has been facilitated by the inclusion of propeptides between the signal peptide and the amino acid sequence of the actual exogenous protein. A signal peptide is a peptide that directs or facilitates secretion of a protein from a cell. These propeptides have been useful for stabilizing the secreted protein against proteolytic activities and enhancing secretion of the protein by providing an endogenous signal peptidase processing site.

The use of propeptides for the secretion of exogenous proteins in *Streptomyces* has been described using signals and propeptides from β-galactosidase for interleukin-1β (Lichenstein, *et al.*, 1988) and thaumatin (Illingworth, *et al.*, 1989); from tendamistat for proinsulin (Koller, *et al.*, 1989), interleukin-2 (Bender, *et al.*, 1990a) and hirudin (Bender, *et al.*, 1990b); and from serine protease inhibitor for domains of immunoglobulin G (Yoshikata, *et al.*, 1993; Ueda, *et al.*, 1993), and CD4 (Fornwald, *et al.*, 1993). In most of the above-cited cases, the removal of the propeptide results in a decrease in the level of secreted exogenous protein.

The processing of the signal peptide leaves the propeptide fused to the amino terminus of the exogenous protein. This situation is undesirable for a therapeutic protein as the presence of the propeptide may alter its biological or immunological activities. To overcome these problems, the propeptide is generally removed by treatment with an exogenous protease(s) following recovery of the secreted protein. Proteases with relatively narrow substrate specificities, e.g., factor Xa, thrombin, collagenase and enterokinase, have been used for the removal of propeptides. Thus a need

exists for the use of proteases with additional substrate specificities to process proteins as desired.

SUMMARY OF THE INVENTION

This invention meets needs in the art by (A) providing new *Streptomyces* strains which lack or have impaired degradative proteases, (B) providing a method for improving the expression and secretion of correctly processed exogenous proteins, (C) providing a method to correctly process proteins with additional sequences incorporated into a genetic construct to prevent degradation by other proteases, (D) inhibiting certain *Streptomyces* proteases, and (E) providing new uses of *Streptomyces* proteases.

A protease isolated and characterized in this invention is a tripeptidyl aminopeptidase designated Tap. Another protease with similar proteolytic activity is a tripeptidyl aminopeptidase, whose nucleotide sequence has a strong homology to the nucleotide sequence of a subtilisin protease and, for that reason, was designated Ssp. Another protease is an X-Pro metalloendoproteinase designated XP-Mep. By this invention, other degradative proteases, each with a characteristic action, were isolated and characterized. This invention includes the nucleic and amino acid sequences, the promoter and signal sequences of such proteases. Protease X of PCT/US92/05532 has different DNA and amino acid sequences than the proteases of this invention and cleaves different substrates than those of this invention.

This invention further includes introducing a DNA sequence encoding a protease into a recombinant vector which, when transformed into suitable host strain, produces an exogenous protease having the biological activity of a wild type protease. Both prokaryotic and eukaryotic hosts may serve as hosts, for producing such proteases.

DNA sequences for a protease of the present invention with a suitable, promotor and/or signal may be used in vectors to direct the expression and secretion of exogenous proteins from *Streptomyces*, or to express increased levels of protein by using multicopy vectors.

The improved strains of this invention have impaired protease production systems, resulting in an increase in quality, quantity or stability of expressed proteins. Such a strain is impaired by deleting, mutating or substituting one or more nucleotides in the sequence encoding for at least one protease. The strains of this invention are produced from the *Streptomyces* genera including species consisting of *S.lividans, S.ambofaciens, S.coelicolor, S.alboniger, S.parvulus* and *S. rimosus*.

A method of fermentation using genetically engineered *Streptomyces* host cells with impaired protease activity is part of this invention. The method includes the steps of: (a) constructing *Streptomyces* host cells with impaired protease activity and which express a desired exogenous protein under suitable conditions; and (b) placing the cells in suitable conditions for expression of the desired protein. The method of fermentation is used to express GM-CSF, IL-3, IL-6, EPO, TNF, SCF, IL-7, IL-2 or any other desired protein.

A method for improving the secretion of proteins from a genetic expression system is part of this invention. The primary amino acid sequence of certain secreted proteins may impose certain physicochemical properties and/or conformational properties which may interfere with the processing of the signal peptide. Secretion of such proteins by a genetic expression system is improved by adding tripeptides (propeptides) to the amino terminal end of the protein which is a precursor to the desired product of the system. The addition is made immediately adjacent to the signal peptidase cleavage

site. The propeptide is removed from the protein by use of a protease such as Tap, Ssp or another protease of this invention.

The inhibitors of this invention are capable of protecting expressed proteins. An inhibitor of this invention is L-alanyl-L-prolyl-L-alanine chloromethyl-ketone, its salts and analogs. This invention includes the use of this inhibitor to inhibit one or more tripeptidyl aminopeptidases derived from *Streptomyces*.

Kits of this invention contain one or more isolated and purified proteases derived from *Streptomyces* to remove one or more amino acids from a protein. An example is a kit for ELISA which consists of a protease and a substrate cleaved by the protease.

DEFINITIONS

In this application, the following terms have the following meanings, unless the context requires otherwise:

"Endogenous protease" means a protease which occurs naturally in a particular host cell and which cleaves one or more of the substrates referred to in this application.

"Exogenous" refers to DNA sequences and proteins which do not occur naturally in a host cell.

"Host cell" means a prokaryotic or eukaryotic cell, strain, species or genera, that may be suitable for introduction and for expression of exogenous DNA sequences.

"Impaired" means reduction or elimination of an activity of a protease produced by a nucleotide sequence compared to the activity of the wild type protease.

"Proteins" includes amino acids, peptides and polypeptides.

"Wild type" means the activity characteristic of a naturally occurring nucleic acid and protein.

"Selective inhibitor" means a molecule that inhibits proteins, such as proteases, on a selective basis.

ABBREVIATIONS

-3	=	protein from which three amino acid residues have been removed from the N-terminus of the protein
-4	=	protein from which four amino acid residues have been removed from the N-terminus of the protein
-6	=	protein from which six amino acid residues have been removed from the N-terminus of the protein
aa	=	amino acid
pNA	=	L-alanine p-nitroanilide
AA-pNA	=	L-alanyl-L-alanine p-nitroanilide
AAPA-pNA	=	L-alanyl-L-alanyl-L-prolyl-L-alanine p-nitroanilide
AMC	=	7-amino-4-methylcoumarin
APACMK	=	L-alanyl-L-prolyl-L-alanine chloromethylketone
APA-AMC	=	L-alanyl-L-prolyl-L-alanine 7-amino-4-methylcoumarin
APF-bNA	=	L-alanyl-L-prolyl-L-phenylalanine beta-naphthylamide
APA-pNA	=	L-alanyl-L-prolyl-L-alanine p-nitroanilide
APM-pNA	=	L-alanyl-L-prolyl-L-methionine p-nitroanilide

APPS-bNA	=	L-alanyl-L-prolyl-L-prolyl-L-serine beta-napthylamide
APS-bNA	=	L-alanyl-L-prolyl-L-serine beta-naphthylamide
bNA	=	beta-naphthylamide
Boc	=	N-t-butoxycarbonyl
Boc-AAPA-pNA	=	N-t-butoxycarbonyl-L-alanyl-L-alanyl-L-prolyl-L-alanine p-nitroanilide
Boc-APARSPA-bNA	=	N-t-butoxycarbonyl-L-alanyl-L-prolyl-L-analyl-L-arginyl-L-seryl-L-prolyl-L-alanine beta-napthylamide
Boc-APS-bNA	=	N-t-butoxycarbonyl-L-alanyl-L-prolyl-L-serine beta-napthylamide
Boc-FSR-AMC	=	N-t-butoxycarbonyl-L-phenylalanyl-L-seryl-L-arginine 7-amino-4-methylcoumarin
Boc-LSTR-pNA	=	N-t-butoxycarbonyl-L-leucyl-L-seryl-L-threonyl-L-arginine p-nitroanilide
D-FPR-bNA	=	D-phenylalanyl-L-prolyl-L-arginine beta-napthylamide
D-PFR-pNA	=	D-prolyl-L-phenylalanyl-L-arginine p-nitroanilide
DMSO	=	dimethyl sulphoxide
EDTA	=	ethylenediaminetetraaceti c acid
ELISA	=	enzyme-linked immunosorbent-assay

FPLC	=	fast protein liquid chromatography
GPL-bNA	=	Glycyl-L-prolyl-L-leucine beta-napthylamide
GP-pNA	=	Glycyl-L-proline p-nitroanilide
GPM	=	Glycyl-L-prolyl-L-methionine
GR-pNA	=	Glycyl-L-arginine p-nitroanilide
HEPES	=	N-2-hydroxyethylpiperazine-N'-2-ethanesulphonic acid
HOHH	=	2-hydroxy-6-oxohepta-2,4-dienoate hydrolase
iPrOH	=	isopropanol
L-pNA	=	L-leucine p-nitroanilide
MNNG	=	N-methyl-N'-nitro-N-nitrosoguanidine
N-(3-carboxypropionyl)-APS-bNA	=	N-(3-carboxypropionyl)-L-alanyl-L-prolyl-L-serine beta-napthylamide
N-Ac	=	N-acetyl
N-Ac-APA-pNA	=	N-acetyl-L-alanyl-L-prolyl-L-alanine p-nitroanilide
N-Ac-APPT-bNA	=	N-acetyl-L-alanyl-L-prolyl-L-prolyl-L-threonine beta-napthylamide
N-Bz	=	N-benzoyl
N-Bz-GSHLV-4MbNA	=	N-benzoyl-L-glycyl-L-seryl-L-histidinyl-L-leucyl-L-valine 4-methoxy-beta-napthylamide

N-Bz-R-pNA	=	N-benzoyl-L-arginine p-nitroanilide
N-Bz-VGR-pNA	=	N-benzoyl-L-valyl-glycyl-L-arginine p-nitroanilide
N-Bz-VLK-pNA	=	N-benzoyl-L-valyl-L-leucyl-L-lysine p-nitroanilide
nt	=	nucleotide
ORF	=	open reading frame
PAGE	=	polyacrylamide gel electrophoresis
pEGR-pNA	=	L-pyroglutamyl-glycyl-L-arginine p-nitroanilide
Pd	=	Palladium
PMSF	=	phenylmethanesulfonyl fluoride
pNA	=	p-nitroaniline
P-pNA	=	L-proline p-nitroanilide
R-pNA	=	L-arginine p-nitroanilide
SDS	=	sodium dodecyl sulphate
S-bNA	=	L-serine beta-napthylamide
SPA-bNA	=	L-seryl-L-prolyl-L-alanine beta-napthylamide
Ssp	=	*Streptomyces* Subtilisin-like protein
ssp	=	gene encoding Ssp
Tap	=	tripeptidyl aminopeptidase-S
tap	=	gene encoding Tap
TSB	=	Trypticase Soya Broth
XP-Mep	=	X-Pro Metalloendoproteinase

DESCRIPTION OF DRAWINGS

FIG. 1. Degradation of GM-CSF and IL-3 by *S.lividans* fermentation broth (gel electrophoresis).

FIG. 2. Purified Tap analyzed by SDS-PAGE.

FIG. 3. Inhibition of Tap by PMSF: IL-3 assay.

FIG. 4. Cleavage of synthetic substrates by *S.lividans* fermentation broth.

FIG. 5. Nucleic acid and encoded amino acid sequences of *tap*.

FIG. 6. (A) Common restriction map for *tap*-containing plasmid DNA isolated from clone P3-13 and P3-5.
(B) The *tap* deletion clones.
(C) The *tap* integration clones.

FIG. 7. Conversion of GM-CSF to GM-CSF(-3) upon incubation with fermentation culture supernatants from cells carrying the *tap* clones.

FIG. 8. SDS-PAGE stained with Coomassie Brilliant Blue staining of fermentation supernatants from cultures of *S.lividans* 66 and *S.lividans MS7* mutant protoplasts transformed with the GM-CSF expression vector pAPO.GM-CSF.

FIG. 9. Conversion of GM-CSF to GM-CSF(-3) upon incubation with fermentation culture supernatants from cells carrying P5-4, P5-6 and P5-10.

FIG. 10. Restriction map for P5-4 and P5-15 and their deletion clones.

FIG. 11. Nucleic acid and encoded amino acid sequences of P5-4.

FIG. 12. Comparison of the predicted amino acid sequence encoded by P5-4 DNA and that of subtilisin BPN'.

FIG. 13. Restriction map for P5-6 and P5- 17n and their deletion clones.

FIG. 14. Nucleic acid and predicted amino acid sequences of P5-6.

FIG. 15. Comparison of the amino acid sequences of Tap and P5-6.

FIG. 16. Restriction map for P5-10 and its deletion clones.

FIG. 17. Nucleic acid and predicted amino acid sequences of P5-10.

FIG. 18. Restriction map for P8-1 and P8-2 and their deletion clones.

FIG. 19. Nucleic acid and predicted amino acid sequence of P8-2.

FIG. 20. Synthetic DNA sequence encoding SCF.

FIG. 21. Synthetic DNA sequence encoding IL-7.

FIG. 22. Synthetic DNA sequence encoding EPO.

FIG. 23. Demonstration of the use of Tap in ELISA technology by standard calibration curve in hIL-3

FIG. 24. SDS PAGE silver stained showing recombinant stem cell factor (SCP) secreted by AP3, AP6, APO and APz vectors.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

Tap A previously unknown protease, a tripeptidyl aminopeptidase designated Tap, derived from *Streptomyces*, has been identified, isolated, and characterized. The enzyme was purified by pH precipitation and chromatography. The pure protease had an apparent molecular weight of 55 kd as determined by SDS-PAGE.

Ssp Another previously unknown protease, a tripeptidyl aminopeptidase designated Ssp, derived from *Streptomyces*, has been identified, isolated, and characterized. The enzyme was purified by pH precipitation and chromatography. Ssp showed a significant amino acid sequence homology to that of *Bacillus subtilis* protease BPN' and was therefore designated Ssp (Subtilisin-like protein).

XP-Mep Another previously unknown protease, an X-Pro metalloendoproteinase derived from *Streptomyces*, has been identified, enriched and characterized. The enzyme was purified by a combination of chromatography and electrophoresis. The protease has an apparent molecular weight of 57-60 kd as determined by SDS-PAGE.

Other Proteases Other previously unknown proteases derived from *Streptomyces* have been identified, and partially characterized.

Degradation Products The major degradation products of Tap were isolated and analyzed by amino acid sequencing. This analysis indicated that the major degradation products were produced by the removal of the N-terminal tripeptides, APA and APM, from GM-CSF and IL-3 (FIG. 1A, Lanes 2, 3 and FIG. 3, lanes 5-8 respectively).

Based upon this information, the molecule APA-pNA and other analogues were synthesized as a potential substrate for Tap. This and several commercial substrates were employed in a survey of proteolytic activities in *S.lividans* fermentation broths (FIG. 4).

APA-pNA cleaving activity was greater than any other activity measured in the broth. This data suggested that one or more tripeptidyl proteases were responsible for the activity. Additionally, the lack of this activity towards the amino-blocked analog, N-Ac-APA-pNA, indicated that the proteases responsible were one or more aminopeptidases.

The presence of a two step cleavage via removal of AP followed by hydrolysis of A-pNA was eliminated by the low activities towards AP-pNA and A-pNA. This presence of an elastase-like endoproteinase was discounted by the lack of measurable cleavage of Boc-AAPA-pNA or N-Ac-APA-pNA in the filtrates (Ashe, *et al.*, 1977; Hunkapiller, *et al.*, 1976; Fink, *et al.*, 1976).

Pure Tap cleaved the N-terminal tripeptide from GM-CSF and cleaved the N-terminal tripeptide from IL-3. When GM-CSF or IL-3 were used as a substrate, the cleaved products produced by pure Tap were identical to the major degradation products found in *Streptomyces* fermentations.

Pure Ssp cleaved the N-terminal tripeptide from M-CSF GM-CSF and sequentially removed tripeptide units from the N-terminus of IL-3. Unlike Tap, Ssp showed no preference for APA-pNA relative to APM-pNA.

The proteolytic activity of XP-Mep was followed by assaying the conversion of recombinant M-CSF GM-CSF to the N-terminally degraded forms, M-CSF GM-CSF(-4) and M-CSF GM-CSF(-6) (Fig. 1A, Lane 4). When M-CSF GM-CSF was secreted by *S.lividans*, these degradation products were found in the fermentation broth.

Studies with IL-3 and IL-6, indicated that XP-Mep was also capable of cleaving these proteins. SDS-PAGE analysis revealed that the pattern of IL-3 degradation generated by XP-Mep was very similar to that obtained during fermentation. The endoproteolytic cleavages generated by XP-Mep significantly reduced the

yield of protein biopharmaceuticals obtained from fermentation with *S.lividans*. The losses were particularly severe when the enzyme produced synergistic degradation with Tap and Ssp by removing N-terminal segments resistant to these enzymes.

Inhibitor A selective inhibitor of the proteases Tap and Ssp, L-alanyl-L-prolyl-L-alanine chloromethylketone (APACMK), was designed, synthesized, and applied to inhibit these proteases. APACMK stopped the release of p-nitroaniline from APApNA by Tap and Ssp. APACMK stopped the cleavage of M-CSF GM-CSF by Tap and Ssp. In fermentations of M-CSF GM-CSF, APACMK prevented cleavage of GM-CSF by Tap and Ssp during fermentation but did not significantly retard the rate of cell growth.

Impaired Strain DNA sequences adjacent to the *tap* and *ssp* genes were used to construct a subclone in which the *tap* and *ssp* genes were precisely deleted. This deletion clone was then substituted into the chromosomes of *S.lividans* 66 strains by homologous recombination to replace the wild type *tap* and ssp loci with mutant *tap* and *ssp* genes.

Disruption of the chromosomal *tap* gene in *S.lividans* resulted in a reduction in the generation of -3 forms of at least tenfold, indicating that this enzyme was responsible for the majority of the activity observed in *S.lividans* strains. Deletional inactivation of the *ssp* gene resulted in a further reduction in the ability of cell-free broth to hydrolyse the chromogenic substrate APA-bNA. Strains carrying chromosomal DNA deletions such as those described herein generally exhibited significantly lower protease activity, reducing the degradation of proteins expressed by genetically engineered host cells, and enabling higher recovery of secreted proteins from the culture supernatant produced by fermentation of the host strain in liquid medium.

Methods of preparing nucleic acid sequences capable of coding for the impaired proteases include: site specific mutagenesis to alter the sequence coding for an essential component for the activity or the expression of the protease; and deletion or mutation of the wild type gene by exposure to mutagens. Generally, the insertion of an impaired gene together with the deletion of a wild type gene is preferred.

Recombinant vectors are useful both as a means for preparing quantities of the protease-encoding DNA itself, or as a means of producing defective proteases to transform recombinant host cells for use in fermentation processes to produce various proteins.

Uses of the Proteases The proteases of this invention are used: (1) in processing recombinant proteins by removing propeptides or fused polypeptides incorporated into a genetic construct, thereby to facilitate expression and secretion to prevent degradation by other proteases, to facilitate purification of polypeptides, etc.; (2) in removing blocked N-terminals from proteins; (3) in digesting connective tissue proteins; (4) in effecting an immunoassay to enable the removal of one or more amino acids from a protein; (5) in carrying out a coupled assay for aminopeptidases to remove peptide blocking groups.

Protease XP-Mep of this invention may be used in a coupled assay for aminopeptidases, such as Tap or dipeptidyl peptidase IV, by removing a peptide blocking group thereby revealing a free amino terminal with an X-Pro sequence for reaction with the targeted enzyme.

Because XP-Mep is capable of digesting gelatin and collagen-like sequences, it is useful in processes involving digestion of connective tissue proteins. Such processes may include a) reduction of scar tissue in muscles, ligaments and tendons, b) wound debridement, c) meat tenderization, and d) cosmetic skin treatment. The

protease XP-Mep is useful as a scientific tool for the removal of blocked N-terminals from proteins to facilitate protein sequencing or for the digestion of proteins for peptide mapping and sequencing studies. The protease XP-Mep is also useful for the processing of recombinant proteins by removing propeptides or fusion proteins incorporated into a genetic construct to prevent degradation by other proteases, to enhance the level of expression of the desired protein, or to facilitate purification of a desired protein.

In addition, to uses previously described, the protease Tap of the present invention is used in coupled assays for proteases, especially those which require an extended P' peptide sequence for substrate recognition, in which the target enzyme cleaves the substrate generating a free amino terminal 3 (or multiples of 3) residues from a reporter group, such as p-nitroaniline, which is then released by Tap yielding a readily measurable and quantifiable colorimetric or fluorometric change.

The protease Tap of this invention is useful in protein sequencing for the removal of proline-containing tripeptides from proteins and peptides thereby giving cleaner sequence data by eliminating residual signals caused by the incomplete cyclization of proline during sequence cycles.

Kit Kits of this invention contain isolated and purified proteases derived from *Streptomyces* to remove one or more amino acids from a protein. The uses of high performance immunoassay have increased greatly in the last decade, extending to almost every discipline in the life sciences. In the majority of applications, antibodies are labelled with enzymes, biotin or fluorochromes, and serve as components of a signal generating/amplifying system. This technology has a broad applicability and can be used in a wide variety of

laboratory techniques including enzyme-linked immunosorbent-assay (ELISA), immunoblotting, immunohisto/cytochemistry and immuno-electrophoresis. Example 31 shows how one can use Tap in the most widely used technique - microwell ELISA. Ssp or another protease of this invention is also suitable for this use.

A kit for ELISA would consist of:

(1) A protease covalently linked to biotin or other carrier capable of participating in the formation of an antigen-antibody complex (example: Tap or Ssp covalently linked to a goat antirabbit IgG);

(2) A substrate, APA-pNA or APA-AMC, which is cleaved by the protease bound in the antigen-antibody complex thereby generating an increase in light absorbance at 405nm with APA-pNA as substrate or an increase in fluorescence when an excitation/emission near 380/460nm is employed with APA-AMC as substrate.

There would be some advantages to using Tap in the ELISA system compared to the common enzymes. The substrates for Tap are more stable, less sensitive to interference due to presence of excipients, and more reliable. The reaction could be incubated much longer and could be measured at any time without stopping the reaction. If necessary, the reaction could be stopped specifically by APA-CMK. Tap activity would not be affected by peroxidases, catalases, phosphatases, chelators, or sodium azide which may interfere with common ELISA enzymes. Using Tap in ELISA would not compromise the sensitivity and may even increase sensitivity by using fluorescent substrates.

Example 1. Purification of Wild-Type Tripeptidyl Aminopeptidase

S.lividans 66 was grown in 11 liters of minimal media (minimal media = 12g Difco Soytone, 10.6g K_2HPO_4, 5.3g KH_2PO_4, 2.5g $(NH_4)_2SO_4$, and 1.0g $MgSO_4$-$7H_2O$ per liter) for 24 hours at 32°C with stirring at 300 rpm in a Chemap

fermenter. Cells were removed from the media by ultrafiltration with a 0.45μm filter (Pellicon System, Millipore). Proteins in the filtrate were concentrated 20 fold by ultrafiltration employing a membrane with a 10 kd cutoff (Millipore). The protease activity was followed by assaying with APApNA and M-CSF GM-CSF as described in Example 2. The protease was precipitated at 4°C by lowering the pH to 4.0 with 0.1M HCl. The precipitate was collected by centrifugation (Model J2-21, Beckman) at 10,000g at 4-10°C and was re-dissolved in 50 ml 10mM Tris-HCl, pH 8.0. After dialysis against 4 liters of the Tris buffer at 4°C, the protease was loaded at ambient temperature onto a 1.6 x 10cm anion exchange column (Q-Sepharose Fast Flow, Pharmacia) equilibrated with the Tris buffer. After washing with equilibration buffer, the bound protease was eluted with a 200ml gradient from 0 to 500mM NaCl at a flow rate of 2ml/minute. The active fractions were pooled to a concentration of 2M in ammonium sulfate. This material was loaded at ambient temperature onto a 1.6 x 10cm hydrophobic interaction column (Phenyl-Sepharose Fast Flow, Pharmacia) equilibrated in 10mM Tris-HCl, pH 8.0, 2M ammonium sulfate. After washing with equilibration buffer, the column was eluted with a 200ml gradient from 2 to 0M ammonium sulfate at a flow rate of 2ml/minute. The active fractions were assayed for purity by SDS-PAGE. The purification of Tap is summarized in FIG. 2 and shows about 20 μg of protein were denatured under reducing conditions and analyzed by SDS-PAGE on 10% polyacrylamide gel. (A) represents purification of wild type. st=Molecular weight standards; Lane 1=Broth obtained after cell removal and concentration of broth by ultrafiltration through a 10kd membrane; Lane 2= Re-dissolved precipitate pH 4.0; Lane 3= Q-Sepharose chromatography pool; Lane 4=Phenyl-Sepharose chromatography pool. (B) represents purified Tap from

the universal search for expressor strain (see Example 12). st=Molecular weight standards; Lane 1=Tap purified from fermentation of the expressor (P3-5) strain.

Example 2. Assays of Tap Activity

Aliquots of Tap column fractions were diluted 100-fold with 20mM Tris-HCl, pH 8.0.

M-CSF GM-CSF as Substrate

To 10μl of rhM-CSF GM-CSF (1μg/μl, Cangene) and 20μl 20mM Tris-HCl, pH 8.0, 20μl of Tap were added. The samples were incubated at 37°C for 2 hours. 20μl of 125mM Tris-HCl, pH 6.8, 0.1% bromophenol blue in 50% aqueous glycerol were added. Products were separated by native gel electrophoresis at constant current on a 17% polyacrylamide gel by a modification of the method of Davies (Davies, 1964) in which the pH of all buffers was modified with H_2SO_4. Products were visualized by staining with Coomassie Blue G-250.

FIG. 1(A) demonstrates this technique by analyzing the M-CSF GM-CSF degradation, where Lane 1 shows intact, full length M-CSF GM-CSF. Lane 2 shows M-CSF GM-CSF from *S.lividans* fermentation. Lane 3 shows degraded isolated M-CSF GM-CSF (-3). Lane 4 shows a mixture of degraded isolated M-CSF GM-CSF(-4) and M-CSF GM-CSF(-6.)

IL-3 as Substrate

To 50μl 20mM Tris-HCl, pH 8.0, 40μl rhIL-3 (2.5μg/μl, Cangene) were added followed by 10μl Tap. The samples were incubated at 37°C. 25μl aliquots were withdrawn at the desired time points and frozen on crushed dry ice. The products were separated by isoelectric focusing from pH 3-10 using Pharmalyte 3-10 (Pharmacia) ampholytes (FIG. 3 - see Example 3). Products were visualized by staining with Coomassie Blue G-250. Intact IL-3 had a pI= 7.4. IL-3(-3) demonstrated a pI= 7.1.

APA-pNA as Substrate

The assay was conducted in a 96 well microtiter plate. To each well in the assay, 50µl 100mM Tris-HCl, pH 8.0, were added followed by 25µl 3.2mM APA-pNA. 25µl of Tap were added to the wells and the absorbance was read at 405nm. The assays were incubated at 37°C for 2 hours. The absorbance was read at 405nm. The activity (release of p-nitroaniline) was calculated from the change in absorbance.

A Survey of Proteolytic Activity in *S.lividans* Fermentation Broths

FIGURE 4 shows the quantification of proteolytic activities in the fermentation broth as measured with synthetic substrates. The assays were conducted in 50mM Tris-HCl, pH 8.0 with 0.8mM substrate incubated at 37°C. The change in absorbance at 405nm was measured after 1, 2, and 4 hours of incubation. The results were reported as micromoles of p-nitroaniline released in 1 hour by 1.0 ml of fermentation broth. 1= APA-pNA; 2= D-PFR-pNA; 3= L-pNA; 4= R-pNA; 5= P-pNA; 6= AP-pNA; 7= A-pNA; 8= AA-pNA; 9= N-Benzoyl-R-pNA; 10= Boc-AAPA-pNA; 11= N-Acetyl-APA-pNA; 12= N-Benzoyl-Y-pNA.

Other Assays and Characterizations

The cleavage specificity of the pure enzyme was examined with both proteins and synthetic substrates. The enzyme readily removed the first N-terminal tripeptide from M-CSF GM-CSF and IL-3 but failed to cleave IL-6 (possessing an additional N-terminal alanine residue). Attempts to force cleavage of IL-6 and further cleavage of M-CSF GM-CSF and IL-3 by increasing the enzyme: substrate ratio by a factor of 1,000 while simultaneously extending the digestion time by a factor of 10 yielded no new products. This resistance may be a function of the primary sequence or protection of the cleavage site by the folding of the target molecule. IL-3 underwent a significant shift in pI from 7.4 to 7.1 upon removal of the APM tripeptide

suggesting a structural rearrangement or the possible formation of a salt bridge by the new N-terminal.

Crystal structures showed the first 14 residues of M-CSF GM-CSF were solvent accessible and the M-CSF GM-CSF(-3) form could be further degraded by enzymes attacking the N-terminus while NMR studies indicated that the N-terminus of IL-6 was also accessible (Diedrichs, *et al.*, 1991; Proudfoot, *et al.*, 1993). It is also possible that the enzyme did not cleave tripeptides containing charged residues. EPO(-3) and (M-CSF GM-CSF(-3) have arginine as their N-terminal residues and appeared to resist further cleavage. The resistance of IL-6 may reside in the primary sequence, Ala-Pro-Val-Pro-Pro. The bulky valine side chain may prevent proper alignment in the active site. Alternatively, the secondary amide may be resistant to cleavage.

APA-pNA, APA-AMC, APM-pNA, and APS-bNA were synthesized by coupling Boc-AP (Bachem) with A-pNA (Sigma), A-AMC (Bachem), M-pNA (Bachem), and S-bNA (Bachem), respectively, by the mixed anhydride method followed by de-blocking with trifluoroacetic acid. AAPA-pNA was prepared by de-blocking Boc-AAPA-pNA (Bachem) with trifluoroacetic acid. N-Carbobenzoxy-APARSPA-pNA was synthesized by manual solid phase methods employing a SASRIN resin (Bachem).

The protease was active against APA-pNA and the fluorogenic analog APA 7-amido-4-methylcoumarin (APA-AMC). The kinetic constants for cleavage of APA-pNA by TAP were Km = 37μM and V_{max} = 55 μmoles $min^{-1} mg^{1}$ at pH 8.0. The enzyme cleaved APA-pNA 10 times as rapidly as the IL-3 analog, APM-pNA. This result agreed with those obtained when comparing the rates of M-CSF GM-CSF and IL-3 cleavage. The tripeptide beta-naphthylamide (bNA), APS-bNA, was also hydrolyzed.

The elastase substrate, N-Ac-APA-pNA, was completely resistant to cleavage by Tap. The N-blocked Boc-APS-bNA was not hydrolyzed by Tap. The extended N-blocked M-CSF GM-CSF analogs, Boc-APARSPA-bNA and N-carbobenzoxy-APARSPA-pNA, were not affected by Tap, but in a coupled assay the reporter groups were rapidly released by simultaneously incubating the molecules with Tap and TPCK-treated trypsin (Sigma). This clearly demonstrated absolute requirement for an un-blocked N-terminal group in the substrate. Additionally, Tap did not release the reporter group from Boc-AAPA-pNA (Bachem), N-Bz-R-pNA (Sigma), N-Bz-VGR-pNA (Sigma), A-pNA (Sigma), R-pNA (Sigma), L-pNA (Sigma), P-pNA (Sigma), S-bNA (Bachem), AA-pNA (Sigma), GP-pNA (Sigma), D-PFR-pNA (KabiVitrum), or AAPA-pNA (Cangene Corporation). The enzyme has shown no ability to act on monoamino acid, dipeptide, or tetrapeptide substrates. The lack of activity towards D-PFR-pNA may result from the N-terminal D-isomer or the presence of the charged arginine as discussed herein.

The effect of pH on the activity of Tap has been examined. When APA-pNA was used as a substrate, the enzyme was active from between pH 5.0 and 9.5 with the maximal activity obtained from between 8.0 and 8.5.

The enzyme cleaved M-CSF GM-CSF from between pH 4.0 and 10.0 with greatest activity from between 5.0 and 9.0. The broad maximum for M-CSF GM-CSF reflected the high sensitivity of this substrate to Tap. The enzyme cleaved IL-3 from between pH 5.0 and 9.0 with maximal activity attained between 7.0 and 8.5. The enzyme was similar in size to the lysosomal tripeptidyl aminopeptidases isolated from bovine anterior pituitary glands and porcine ovaries but does not exhibit the acidic pH optimum of those serine proteases (McDonald, *et al.*, 1985; Doebber, *et al.*, 1978). The pH optimum more closely resembled that of the extracellular enzymes from human liver and erythrocytes but was considerably smaller than

these 135,000 dalton enzymes (Balow, *et al.*, 1983). Examination of the *tap* DNA sequence may provide some insight into evolutionary relationships between the prokaryotic and eukaryotic enzymes.

Example 3. Inhibitors of Tap

Inhibitor Survey: Assay With APA-pNA

Tap was not inhibited or stimulated by chelating agents, divalent cations, or sulfhydryl reagents, thereby eliminating any relation to the subtilisins, thiol proteinases, or metalloproteinases. Pepstatin was also ineffective. The failure of the peptide aldehydes, elastatinal and chymostatin, may reflect Tap's lack of endoproteolytic activity.

An inhibitor survey indicated that Tap was a serine protease. Table I shows the inhibition of Tap activity by various protease inhibitors. The protease and inhibitor were preincubated for 15 minutes at 22°C. Substrate was added and the mixture was incubated at 37°C. Activity was measured by monitoring the change in absorbance at λ= 405nm.

The enzyme was inhibited by the serine protease inhibitor, phenylmethanesulfonyl fluoride (PMSF). Treatment of Tap with PMSF inhibited cleavage of M-CSF GM-CSF, IL-3, and APA-PNA.

Table I

Inhibition of TAP in the APA-pNA Assay

Sample	Concentration	Residual Activity
Enzyme only	------	100
PMSF	1.6 mM	7
$HgCl_2$	0.1 mM	99
	1.0 mM	93
$CaCl_2$	1.0 mM	96
	10 mM	97
$CoCl_2$	1.0 mM	98
	10 mM	97
EDTA	1.0 mM	95
	10 mM	95
IDA	1.0 mM	82

DTT	1 mM	86
DTT + EDTA	1 mM + 10 mM (respectively)	97
Elastatinal	0.1 mM	97
Chymostatin	0.1 mM	98
Pepstatin	0.1 mM	95
Benzamidine	10 mM	94

Inactivation of Tap with PMSF: Assayed with IL-3

Tap stock (Example 1) was diluted 100-fold with 20mM Tris-HCl, pH 8.0. A fresh solution of 80mM PMSF was prepared in isopropanol (iPrOH). A stock buffer of 20mM Tris-HCl, pH 8.0 was prepared. Four solutions were preincubated as follows.

iPrOH	= 58µl Stock Buffer + 2µl iPrOH
PMSF	= 58µl Stock Buffer + 2µl PMSF/iPrOH
Tap + iPrOH	= 18µl Stock Buffer + 40µl Tap 2µl iPrOH
Tap + PMSF	= 18µl Stock Buffer + 40µl Tap + 2µl PMSF/iPrOH

These solutions were incubated at 22°C for 30 minutes. When the preincubation was complete, 40µl rhIL-3 (2.5µg/µl, Cangene) were added and incubation was initiated at 37°C. Aliquots of 25µl were removed at 0, 1, 2, and 4 hours. These aliquots were immediately frozen on dry ice. When the sampling process was complete, the products were analyzed by isoelectric focusing from pH 3-10 (Example 2).

FIGURE 3 shows the inactivation of Tap against IL-3. Lanes 1-4 show the incubation of IL-3 with Tap that was treated with PMSF. Lane 1= 4 hours; Lane 2= 2 hours; Lane 3= 1 hour.; Lane 4= 0 hours. Lanes 5-8 show the incubation of IL-3 with uninhibited Tap. Lane 5= 4 hours.; Lane 6= 2 hours.; Lane 7= 1 hour.; Lane 8= 0 hours. Lane 9 is a human carbonic anhydrase marker, pI= 7.4. Lane 10 contains pI markers. As can be seen in Lanes 5-8 of FIG. 3, the IL-3 (pI= 7.4) was completely converted to IL-3(-3) (pI= 7.1) by Tap within 2 hours. Lanes 1-4 show that with PMSF treatment, intact IL-3 is clearly detected after 4 hours. Tap was not affected by sulfhydryl reagents, chelators or aspartyl protease inhibitors (Table I).

Example 4. Amino Acid Sequencing of Tap

Tap was purified as described in Example 1 and was desalted by size exclusion chromatography. An Immobilon polyvinylidenedifluoride (PVDF) membrane (Millipore) was solvated according to the manufacturer's instructions. Tap was adsorbed to the membrane by filtration employing a slot blot assembly. Protein bound to the membrane was visualized with Amido Black. The sample was excised and subjected to automated Edman degradation for 15 cycles.

The N-terminal sequence of the isolated wild-type Tap was NH_2-Asp-Gly-His-Gly-His^5-Gly-Arg-Ser-Trp-Asp^{10}-Arg-Glu-Ala-Arg-Gly^{15}.

Table II

	Mole percentage	
Amino Acid	Protein	DNA
Asp + Asn	13.6	12.4
Glu + Gln	10.9	7.6
Ser	4.7	4.7
Gly	10.0	8.9
His	2.2	2.3
Arg	7.4	7.4
Thr	6.3	6.3
Ala	14.3	14.3
Pro	7.2	7.2
Tyr	3.9	3.8
Val	6.4	7.6
Met	1.2	1.3
Ile	2.3	3.0
Leu	5.6	6.3
Phe	1.7	2.5
Lys	2.5	4.4

The first 15 residues of the N-terminal of the isolated wild-type protease were determined and identically matched amino acids 40-54 derived from the DNA sequence (FIG. 5). Residues -39 to -4 appeared to be a signal peptide. An autolytic tripeptide cleavage removing APA after signal peptide removal would yield the N-terminal found for the secreted protease.

Table II shows the amino acid composition of the wild-type Tap. The amino acid composition derived from the corresponding portion of the *tap* gene DNA sequence (FIG. 5) is shown for comparison. The small differences in composition may be attributable to low level impurities in the enzyme sample.

Example 5. Synthesis of APACMK

21.3g (70mmol) Boc-Ala-Pro (Bachem Biosciences) dissolved in 175ml anhydrous dimethylformamide (DMF) were activated by adding 7.8ml (70.7mmol) 4-methylmorpholine followed by 9.3ml (70.7mmol) isobutylchloroformate at -20°C with stirring. After 15 minutes, 15.1g A-OBz in 175ml anhydrous DMF were added. The solution was stirred for 1 hour at -20°C and then for 17 hours at ambient temperature. The DMF was removed by vacuum rotary evaporation. The residue was taken up in 175ml ethyl acetate and extracted each with 5% citric acid, saturated sodium bicarbonate, water, and brine. The organic layer was dried over anhydrous sodium sulfate for 1 hour. The sodium sulfate was removed by filtration.

2.5g 5% Pd on activated carbon were added and the suspension was agitated under a hydrogen atmosphere for 2 hours. At that time, the starting material had been completely converted to product. The hydrogenation catalyst was removed by filtration through Celite. The solvent was removed by vacuum rotary evaporation.

The resulting 23.7g (66.3mmol) of Boc-APA were dissolved in 140ml anhydrous ethyl acetate and reacted with 7.8ml (70mmol) of 4-methylmorpholine followed by 9.2ml (70 mmol) of isobutylchloroformate at -20°C with stirring. After 15 minutes, a solution of diazomethane in anhydrous ether prepared from 100mmol N-methyl-N-nitroso-p-toluenesulfonamide (Aldrich) was added. After 1 hour at ambient temperature, the solution was extracted twice with 140 ml portions of water. The organic layer was dried over 2 g anhydrous sodium sulfate powder for 1 hour. The solution was removed by decantation. Deblocking of the N-terminal and generation of the chloromethylketone group were achieved simultaneously by adding 100ml of HCl(g) saturated ethyl acetate. The resulting solution was allowed to stand at ambient

temperature for 30 minutes. The product was removed from the organic solvent by extraction into 400ml of water. The aqueous pool was frozen and lyophilized to yield the product, APACMK, as its hydrochloride salt.

Example 6. Inactivation of Tap by APACMK: Assayed with APA-pNA

A stock solution of 10nM Tap in 100mM Tris-HCl, pH 8.0 was prepared. Serial dilutions of 210μM, 21μM, 2.1μM, 210nM, 21nM, and 2.1nM APACMK (Example 5) were prepared. To the microtiter well, 25μl of Tap followed by 25μl of an APACMK dilution or distilled water, for an uninhibited control, were added. The assays were incubated for 20 minutes at 22°C. 50μl 1.6mM APA-pNA were added to each well. The absorbance was read at 405nm then incubated at 37°C. The change in absorbance at 405nm was read after 15 and 60 minutes of incubation.

Example 7. Inactivation of Tap by APACMK: Assayed with M-CSF GM-CSF

A stock solution of 10nM Tap in 20mM Tris-HCl, pH 8.0 was prepared. Serial dilutions of 210uM, 21uM, 2.1uM, 210nM, 21nM, and 2.1nM APACMK (Example 5) were prepared. To 20μl Tap, 20μl of an APACMK dilution (or water for an uninhibited enzyme control) were added and incubated at 22°C for 30 minutes. 10μl of M-CSF GM-CSF (1μg/μl, Cangene) were added and incubated at 37°C for 2 hours. Products were analyzed by native gel electrophoresis as described in Example 2.

Example 8. Inactivation of Tap by APACMK - Determination of Kinetic Constants

A stock solution of 1.1μM Tap in 50mM Tris-HCl, pH 8.0 was prepared. APACMK stock solutions of 11μM, 13.8μM, 17.3μM, 21.7μM, 27.0μM, 54.0μM, 108μM, and 1.08mM were prepared. The substrate solution was 50mM Tris-HCl, pH 8.0, 0.8mM APA-pNA. The inactivation was performed by placing 90μl of Tap (1 nanomole) in a 1.5 ml Eppendorf tube on ice and adding 10μl of water (uninhibited control) or 10μl of APACMK. A 10μl aliquot was removed immediately and was assayed by adding it to a cuvette containing 390μl substrate solution at 22°C. The initial velocity was obtained from the change in absorbance at 405nm during the first 10 seconds of the assay. Additional aliquots were removed at time points and assayed by the same method. At APACMK concentrations greater than 5.0μM in the incubation, it was not possible to remove an aliquot from the incubation before 90% inactivation occurred.

The inhibitor APACMK yielded K_i= 3.3 μM and k_{inact}= 0.14 min^{-1} with >99% inactivation within 6 minutes at 0°C at an inhibitor concentration of 2.7μM and an inhibitor/enzyme molar ratio of 2.7.

Example 9. Application of APACMK in Fermentation

A 100ml volume of media was inoculated in 500ml baffle-bottom flasks with 100μl of *S.lividans* 66 working seed bank material. The cultures were grown in a New Brunswick gyratory incubator at 32°C and 240rpm. The cultures were sampled at 25, 27, 29, 31 and 48 hours post-inoculation and were analyzed by native gel electrophoresis. Following removal of the 25 hour sample, 100mM APACMK in sterile water were added to yield a final concentration of 10μM. A control flask without APACMK was retained. The addition of APACMK

significantly reduced formation of M-CSF GM-CSF(-3) but did not inhibit cell growth.

Example 10. Construction and Screening of a *S.lividans* Genomic Library

A *S.lividans* 66 (Hopwood, *et al*,, 1983) genomic library was made using size fractionated (3-12 Kb) fragments of chromosomal DNA partially digested with *Sau3AI* and ligated into the *Bam*HI site of the bifunctional cloning vector, pSS12 (Butler, *et al.*, 1992). The ligated DNA was used to transform competent cells of *E.coli* HB101 and pooled plasmid DNA was isolated from a mixture of approximately 30,000 transformed colonies grown in SOB medium (Maniatis, *et al.*, 1982) containing ampicillin (Sigma). This DNA was used for transformation of *S.lividans* 66 protoplasts yielding 15,000 transformant colonies resistant to thiostrepton (E.R. Squibb). Two days later the colonies were screened by overlaying with substrate mixture [containing 5 ml phosphate buffer (50 mM, pH 7.0), 25µl GPL-bNA (20 mg/ml in DMSO), 0.1 ml Fast Garnet GBC (10 mg/ml in water)]. The plates were incubated for three minutes at room temperature and washed three times with saline solution (Atlan, *et al.*, 1989, Alvarez, *et al.*, 1985). Positive colonies stained intensely red against a background for pale red colonies.

Two colonies reproducibly showed strong color. Plasmid DNA was isolated from each of these two colonies and the phenotype was retained when the DNA was used to transform protoplasts of *S.lividans* 66. The plasmid DNA from each of these clones (P3-5 and P3-13) was investigated by restriction enzyme analysis. The data indicated that P3-5 and P3-13 were identical (presumably siblings). The peptidase gene was localized within the cloned DNA fragment by monitoring the GPL-bNA hydrolyzing activity of strains containing

various subclones and deletion clones derived from the original clones.

FIGURE 6 (A) shows a restriction enzyme site map of cloned P3-5 DNA. Phenotype in the GPL-bNA hydrolysis agar plate assay is shown qualitatively as the number of + signs judging red color developed on the colonies. The *Eco*RI site shown in parentheses was present in the pSS12 vector adjacent to the *Bam*HI cloning site. None of the three deletion clones, shown in FIG. 6(B), produced any more red color in colonies than did the pSS12 control plasmid Each were scored as "+" due to the background level of hydrolysis from the chromosomal *tap* gene in the *S.lividans* 66 host.

Example 11. Tap Activity of *S.lividans* 66 Strains Carrying the P3-5 and P3-13 Clones

The *S.lividans* 66 strains carrying the P3-5 and P3-13 clone or pSS12 were grown in TSB (containing 1% glucose, 0.1 M MOPS and 20 μg ml^{-1} thiostrepton). Aliquots (40 ml) of each culture were removed at 23 and 29 hours, and the supernatant and mycelium fractions were separated by centrifugation. Aliquots of the supernatant fractions were added to reactions (100 μl) containing various tripeptide-bNA substrates (8nmol) in microtiter wells. After incubation at 37°C for 4 hours, a solution (50 μl) containing Fast Garnet GBC dye was added and the A_{540} was measured in a microtiter plate reader. The results are shown in Table III.

Table III

Tripeptidyl Aminopeptidase Activity
(A_{540} above background)

Sample	GPL-bNA	GPM-bNA	APF-bNA	D-FPR-bNA
Supernatants				
P3-5/23 - Hours	Max	Max	Max	0.02
P3-5/29 - Hours	Max	Max	Max	0.08
SS12/23 - Hours	0.19	0.28	0.63	0.02
SS12/29 - Hours	1.38	2.46	Max	0.17

("Max" indicates a A_{540} reading of >3.0)

At as early as 23 hours of culture, a 1μl aliquot of the supernatant from *S.lividans* carrying the P3-5 clone was showing strong activity against the GPL-, GPM- and APF-bNA substrates. At the same time point, a 25 μl aliquot of the control culture had at least 15 to 20 fold lower activity with the same substrates. However, against the D-FPR- and APF-bNA substrates, the Tap expressor had little increased activity over the control. An aliquot (1 μl) of each supernatant (which was harvested after 23 hours of growth) was added to a reaction containing 4 μg of purified intact M-CSF GM-CSF. Following a 2.5-min. incubation at 37°C, the proteins were analyzed by native PAGE and stained with Coomassie Brilliant Blue.

FIGURE 7 shows the conversion of exogenously added, purified full length M-CSF GM-CSF (lane 1) degraded to GM-CSF(-3) upon incubation with fermentation culture supernatants from culture samples carrying the p3-5 and

P3-13 clones (lanes 2 and 3, respectively). By contrast, no significant degradation was observed when M-CSF GM-CSF was incubated with the control culture (lane 4) due to the small volumes of culture supernatant and short time of incubation used compared to those described in Example 2.

Example 12. Analysis of Extracellular Proteins From *S.lividans* 66 Strains Carrying the p3-5 and p3-13 Clones

The *S.lividans* 66 carrying the P3-5 and P3-13 clones were grown in liquid culture, and supernatant fractions were collected following the teaching of Example 11. As described by Laemmli (1970), samples were prepared from aliquots (200 μl) of the supernatant fractions, and SDS-10% polyacrylamide gels were run at 100 v for 5 to 6 hours. The profile of separated proteins was then visualized by staining with Coomassie Brilliant Blue.

An abundant protein with an apparent molecular weight of 55,000 daltons was present among the extracellular proteins from *S.lividans* 66 carrying either P3-5 or P3-13. From 23 to 29 h of culture, the level of Tap increased to approximately 0.1 mg/ml, relative to the BSA standards that were included on the same gel.

Example 13. Amino-Terminal Sequence of the Tap Protein Purified from an *S.lividans* 66 Strain Carrying the P3-13 Clone

The *S.lividans* 66 strain carrying the P3-13 clone was grown in liquid culture and supernatant fractions were collected, following Example 11. The extracellular proteins were separated by SDS-PAGE, following the teaching of Example 12, and transferred onto Immobilon PVDF (Millipore) membranes as directed by the supplier. After briefly staining the filters with Coomassie Brilliant Blue, the bands containing the major protein were excised from the filter, and subjected to automated Edman degradation analysis. The amino-terminal amino

acid sequence determined was: Asp-Gly-His-Gly-His-Gly-Arg-Ser-Trp-Asp-Arg.

The amino-terminal sequence of the protease from the expressor (P3-5) matched the amino-terminal sequence of the Tap purified from the extracellular broth of untransformed *S.lividans* 66. In addition, the enzymes from both sources had similar substrate specificities and the same apparent molecular weight of 55,000 daltons as determined by SDS-PAGE (FIG. 2). These factors indicate that the protease isolated from wild type *S. lividans* 66 and the P3-5 transformant were the same enzyme.

Example 14. Nucleotide Sequence of the *tap* gene

The nucleotide sequence for the *S.lividans tap* gene is shown in Figure 5. Inspection of the sequence reveals a potential protein encoding region of 537 codons amino acids consistent with the molecular weight of the observed secreted protein of 55kd. The potential protein coding region is indicated by an arrow in FIG. 6(A) and the deduced amino acid sequence is shown in Figure 5 for each codon. The putative translation initiation codon is a TTG which is preceded by a strong potential ribosome binding site GGAGG. The deduced amino acid sequence shows a predicted signal peptide (von Heijne, 1986) which could be cleaved between residues 36 and 37 (ASA↓A). The experimentally determined amino acid sequence of the mature secreted protein was identical to that encoded by residues 41 to 51 of the predicted protein sequence. This suggests that the protein is secreted as a propeptide with a three amino acid propeptide (APA). Since the enzyme removes such tripeptides from polypeptides it is highly likely that the removal of this APA is autocatalytic (although this was not experimentally tested). When the deduced amino acid sequence was compared to sequences in the protein databases, the closest match was with that of the 2-

hydroxy-6-oxo-2,4-heptadienoate hydrolase (HOHH) from *Pseudomonas putida* F1 (Menn, *et al.*, 1991) as follows.

Tap (199) K L N Y L G V S Y G T Y L G A V Y G T L F P D H V R R
M V V (288)

HOHH (98) R V D L V G N S F G G A L S L A F A I R F P H R V R R
L V L (127)

The weak homology includes the GVSYG sequence which includes the serine residue potentially involved in the catalytic action of this protease. The coding sequence is closely followed by a substantial inverted repeat (ΔG=-62kcal.mole^{-1} calculated according to Tinoco, *et al.*, (1973) suggesting that the *tap* gene is transcribed independently of other adjacent DNA sequences. A potential protein encoding region is present upstream of the *tap* gene, however, since an inverted repeat sequence (ΔG=-39kcal.mole^{-1}) follows this region it may be transcriptionally independent of the *tap* gene.

Example 15. Analysis of the Chromosomal *tap* Gene of Various *Streptomyces* Species

Genomic DNA was isolated from the following *Streptomyces* strains. *S.alboniger* 504 (P. Redshaw, Austin College, Texas, USA), *S.ambofaciens* ATCC 23877, *S.coelicolor* M130 (John Innes Institute), *S.fradiae* ATCC 14544, *S.griseus* IMRU 3499, *S.griseus* ATCC 10137, *S.lividans* 66 (B. Pogell, Univ. of Maryland, MD, USA), *S.parvulus* 2283 (John Innes Institute) *S.rimosus* ATCC 10970. 10 μg of each DNA were digested in 100 μl of appropriate buffer for the restriction enzymes *Bam*HI and *Pst*I respectively. 30 units of each enzyme were added together with 1 μl of RNAse A (10 mg/ml, Sigma). The reactions were incubated at 37°C for 3 hours. A further 15 units of enzyme were added and the samples incubated overnight at 37°C. Digestions were terminated by the

addition of 11 μl of stop buffer (0.08% Orange G; 50% glycerol; 67 mM EDTA, pH8). Approximately 3 μg of each digested DNA sample were loaded onto a 1% agarose horizontal gel and electrophoresed at 100V for 4 hours. A molecular weight marker was included (Lambda DNA digested with *Hin*dIII, Bethesda Research Laboratories) to calibrate the gel. After electrophoresis the gel was soaked in 0.25 M HCl, followed by 0.5M NaOH, 1.5M NaCl and rinsed in water. The DNA was transferred to a Nylon membrane (Boehringer Mannheim) using a Vacublot (Pharmacia) apparatus with 20 x SSC buffer for 1 hour at 50 mbars pressure. After transfer the membrane was washed in 2 x SSC and baked for 1.5 hours at 80°C.

The DNA insert fragment from the *Eco*RI site to the right-most *Bam*HI site was isolated by partial *Bam*HI and complete *Eco*RI digestions of the P3-13 DNA. The fragment was subcloned into the *E.coli* plasmid vector pT7T3 (Pharmacia). From this clone it was possible to isolate larger quantities of the same DNA fragment by digestion with *Eco*RI and *Hin*dIII. 0.5 μg of this 3.3 Kb- fragment were labelled according to the manufacturer's recommendations (Boehringer Mannheim) to produce a digoxigenin - labelled probe. 25 ng of probe were used per ml of hybridization solution. Lambda DNA was labelled in the same way to allow visualization of the molecular weight marker fragments. Hybridization was carried out at 68°C overnight using 2.5 ml of hybridization solution per 100 cm^2 of nylon membrane. The hybridization solution contained; 5 x SSC; blocking reagent, 1% (w/v); N-lauroylsarcosine, 0.1% (w/v); sodium dodecyl sulphate, 0.02% (w/v). Filters were prehybridized for 1 hour at 68°C. Probes were boiled for 10 minutes, quick chilled on an ice/NaCl bath, diluted with 100 μl hybridization solution and added to the prehybridized membrane in a stoppered glass bottle. Hybridization and prehybrization were carried out using a Hybaid mini-

hybridization oven. Membranes were washed twice at 68°C for 30 minutes in 5 x SSC, 0.1% SDS (50 ml/100 cm^2 membrane). The membranes were then transferred to plastic containers and processed according to the manufacturer's instructions. Finally, the membranes were transferred to plastic bags, sealed and incubated at 37°C for 30 minutes, and then exposed to X-ray film for 10 minutes. The resulting autoradiogram showed hybridizing bands in all lanes except those containing *S.fradiae* DNA. Identical hybridizing bands were observed with *S.lividans* and *S.coelicolor* with a common band in both *S.griseus* strains as well as the *S.parvulus* DNA. *S.ambofaciens*, *S.rimosus* and *S.alboniger* produced hybridizing bands at different molecular weights, suggesting restriction fragment length differences in these species. This indicates the likely presence of a *tap* gene in other *Streptomyces* species which would be expected to be detrimental to product yield when expression of secreted proteins is desired in these strains. Taken overall the results suggest that the Tap-encoding DNA sequence occurs widely throughout the *Streptomyces* strains examined.

In a similar experiment Southern hybridization analysis of chromosomal DNA, using the plasmid P3-13 as a probe suggested that the DNA contained in P3-13 had not been rearranged during cloning.

Example 16. Construction of A Deletion Subclone From the *tap* Clone

Specific deletions were made in the *tap* clone to localize the gene and enable chromosomal disruption. A 1.2 -Kb DNA fragment was removed between *Bam*HI (1100) and *Bgl*II (2300) (see FIG. 6B) to construct the deletion clone Δ1. P3-5 DNA was digested by means of *Eco*RI and *Bgl*II, and the vector fragment was isolated; and P3-5 was digested with *Eco*RI and *Bam*HI and the 1.1 -Kb insert fragment was isolated. The vector and insert fragments were ligated, using T4 DNA ligase, and used to transform

E.coli. The plasmids were screened by restriction analysis and the correct plasmid, Δ1, used to transform protoplasts of *S.lividans* 66. The *S.lividans* 66 carrying the Δ1 deletion clone was screened with a plate assay using GPL-bNA. A transformant was grown in liquid culture, and the level of Tap activity was determined in a liquid assay using GPL-bNA and APA-bNa substrates. The *S.lividans* 66 carrying the Δ1 deletion subclone had a similar Tap activity to that of the untransformed host strain.

Deletion clone Δ2 was constructed by subcloning the *Eco*RI-*Bgl*II fragment into the vector pSS12 which had previously been digested with *Eco*RI and *Bam*HI. Δ3 was made by digestion of P3-5 DNA with *Bgl*II, followed by religation, resulting in the loss of the 300 nt *Bgl*II fragment around the centre of the *tap* gene. The high level of Tap activity associated with the P3-5 plasmid was not observed with Δ2 or Δ3, confirming that the deletions resulted in loss of enzyme activity.

Example 17. Deletion Clones Used for Integrational Mutation of *tap* into the *S.lividans* 66 Chromosome

This example describes the use of the deletion clones of the *tap* gene for integrational mutation into the *S.lividans* 66 chromosome resulting in inactivation of the wild type *tap* gene. Loss of the wild type *tap* gene occurred by homologous recombination with the integrated mutant DNA sequence using the natural ability of the *S.lividans* host cell to resolve such regions of chromosomal DNA containing directly repeated nucleotide sequences. Resolution occurred apparently at random to produce strains carrying either the wild type parental *tap* gene or the exchanged mutant *tap* gene. Mutant strains were identified by their inability to hydrolyse the chromogenic substrate GPL-bNA.

Subcloning of the DNA insert sequences from the deletion clones was not straightforward due to the presence of multiple *Bam*HI sites. A partial *Bam*HI digestion of P3-5 DNA was followed by a complete *Eco*RI digestion. The 3.1 Kb *tap*-encoding fragment was isolated from an agarose gel and subcloned into the *E.coli* vector pT7T3 which had previously been digested with *Bam*HI and *Eco*RI. Appropriate transformants were identified and the DNA insert was used to create further subclones in the pINT vector as follows: Δ1int was produced by a three way ligation of the *Eco*RI-*Bam*HI, *Bgl*II -*Hind*III (in the polylinker of the pT7T3 vector) fragments from the pT7T3 subclone and the *Eco*RI-*Hind*III fragment produced by digestion of pINT. Δ2int was the result of a direct subcloning of the *Eco*RI-*Bgl*II fragment from the pT7T3 subclone into pINT digested with *Eco*RI and *Bam*HI. Δ3int involved the *Bgl*II-*Hind*III fragment from the pT7T3 subclone and *Bam*HI plus *Hind*III digested pINT. Δ4int was a direct subcloning of the whole inserted fragment in the pT7T3 subclone (*Eco*RI + *Hind*III) into the same sites in pINT. Δ5int was made from Δ4int by digestion with *Bgl*II and religation. The DNA contained within the various Δint clones is shown in FIG. 6C.

Plasmid DNA of the various Δint clones was isolated and used to transform protoplasts of *S.lividans* 66 (wild type) and MS5 (a strain derived from *S.lividans* 66 by deletion of DNA fragments at the *slpA* and *slpC* (Butler, *et al.*, 1992) loci; in addition the *pepP* gene (Butler, *et al.*, 1993) and a second PepP-encoding gene (Butler, *et al.*, 1994) were also subjected to specific chromosomal DNA deletion events, each of which reduced the PepP activity of the *S.lividans* strains). Integrative transformants resistant to thiostrepton were purified and allowed to grow in the absence of thiostrepton to allow recombinational resolution to occur. Strains which had undergone excision events were

easily identified by screening for the loss of the ability to hydrolyse GPL-bNA.

The results obtained were somewhat unexpected. Δ1int did not produce any integrative thiostrepton-resistant transformants from S.lividans MS5 in three independent experiments. Δ2int did lead to integrative transformants, indicating that there was no practical impediment to recombination events at this locus on the *S.lividans* chromosome. Subsequent experiments using Δ1int were successful using *S.lividans* 66 protoplasts (to make a strain designated MS9 which was defective only at the *tap* locus) suggesting that the earlier failure in the MS5 experiment was due to the lower transformation capability of that particular batch of MS5 protoplasts. Δ3int failed to produce integrative transformants, possibly due to the relatively small length of DNA (900nt) available for homologous recombination to occur. However, Δ4int yielded transformants as did Δ5int.

Integrative transformants from *S.lividans* 66 and MS5 using Δ5int were grown in the absence of the thiostrepton selection on agar medium. After sporulation had occurred the spores were harvested and replated onto fresh agar plates. Colonies were screened using the GPL-bNA substrate assay for Tap activity. The frequency of excision events which led to loss of the activity was very low (approximately 1 in 1000). Three colonies were obtained with reduced Tap activity.

Following the teaching of Example 15, chromosomal DNA was isolated from *Streptomyces lividans* 66 and deletion mutant strains and Southern hybridization analysis of the chromosomal *tap* locus was performed. The DNA was digested with *Bam*HI or *Stu*I and transferred to a nylon membrane (Hybond, Amersham). Using a ^{32}P-labelled probe for the *Bgl*II fragment internal to the *tap* gene resulted in a strong band of hybridization at approximately 1.8 Kb in the *Bam*HI digests and two bands in the *Stu*I digests

for both the *S.lividans* control and colony #3 indicating that this DNA fragment was present in both strains. However, no hybridizing bands were observed for colonies 1 and 2 confirming the loss of the 0.3 Kb *Bgl*II fragment. Similar experiments with a 3.3 Kb DNA probe revealed a complex hybridizing band pattern in colony 1 chromosomal DNA whereas colony 2 DNA showed only the expected bands with a reduction in size of one band consistent with the desired specific chromosomal deletion. Colony 2 was designated *Streptomyces lividans* MS7. Another strain was constructed using Δ5int and *S.lividans* 66 protoplasts. This strain was designated MS8 and shown to have properties indistinguishable from those of MS9.

Example 18. The *S.lividans* MS7 Strain Shows a Substantial Reduction in its Ability to Hydrolyse Tripeptide bNA Substrates and GM-CSF *in vitro*

The *S.lividans* MS5 (*tap*$^+$) and MS7 (*tap*$^-$) strains were grown in liquid culture (TSB medium without thiostrepton) and samples of the cultures were collected at various time points during the fermentation. Cell-free supernatant fractions were isolated by centrifugation to remove the mycelial material. Aliquots (50μl) of the supernatants were added to each of the chromogenic tripeptide substrates APA-bNa or GPL-bNA (8nmol) in a final volume of 100μl. After incubation at 37°C for 45 minutes, 50μl of a solution of Fast Garnet GBC dye was added and the A_{540} measured using a microliter plate reader.

The results, summarized in Table IV, indicate that under these assay conditions, the supernatants derived from the MS7 culture were (within experimental error) devoid of any significant hydrolytic ability against these substrates, whereas the supernatant derived from *S.lividans* MS5 showed the ability to rapidly degrade substrates.

Table IV

Tripeptidyl Aminopeptidase Activity of MS5 and MS7

(A_{540})

Substrate	APA-bNA		GPL-bNA	
Strain	MS5	MS7	MS5	MS7
21 hours	1.459	0.337	0.952	0.338
25 hours	2.052	0.456	1.170	0.374
29 hours	1.808	0.390	1.151	0.399
45 hours	1.586	0.382	0.877	0.366

The same supernatant samples harvested at 25 hours were analyzed for the ability to degrade GM-CSF *in vitro* according to the teaching of Example 16. It was clear from native PAGE analysis that the rate of degradation of GM-CSF was significantly slower for MS7 than for MS5.

Example 19. Production of Undegraded GM-CSF by the *S.lividans* MS7 Strain

The GM-CSF expression plasmid vector pAPO.M-CSF GM-CSF was used to transform protoplasts of the *S.lividans* MS7 strain. Following the teaching of Example 11, liquid cultures were prepared from the transformed strain as well as transformants from the *S.lividans* MS5 strain. Cell free broth from the strains was harvested at various time points during fermentation and analyzed by native PAGE (FIG. 8).

The results indicate that while degradation of the secreted GM-CSF occurred in both strains, it was evident in MS7 only at later times of growth, as compared to the MS5 samples. This property of the new *S.lividans* MS7 strain allowed it to be used to produce a higher yield of undegraded GM-CSF than was possible using the wild-type *S.lividans* 66 strain.

Example 20. Production of Mutant Strains of *S.lividans* Defective in Protease Activities Using Chemical Mutagenesis

S.lividans 66 spores were treated with N-methyl-N-nitro-N-nitrosoguanidine (MNNG) (Hopwood, *et al.*, 1985). Briefly, a suspension containing 2.5 x 10^{12} spores in 3 ml of Tris/maleic acid buffer was incubated at 30°C in a preweighed vial containing 10 mg of MNNG (which had been solubilized in 0.5 ml DMSO immediately prior to the addition of the spore suspension). Aliquots of 1 ml were removed from the mixture at 30 minute intervals and washed twice by centrifugation to remove the MNNG. Serial dilutions of the treated spores were plated on agar medium to determine the effectiveness of the mutagenic treatment in terms of the proportion of viable surviving colony forming units remaining compared to untreated spores. Survival rates of 2.8 x 10^{-3} %, 1.2 x 10^{-4} % and 9 x 10^{-6} % were observed after 30, 60 and 90 minutes, respectively.

Two hundred surviving colonies from each of the three treatment times were purified and examined for their ability to grow on minimal medium. Colonies which were unable to grow were classified as auxotrophic mutants of which 1, 4 and 2 were observed at the 30, 60 and 90 minute treatment times, respectively.

Spores from the 60 minute treatment were, therefore, examined for the presence of strains carrying mutations which inactivated specific proteolytic phenotypes. A direct agar plate screening technique was used in which the colonies were overlayed with substrate mixture (containing 0.1 ml of GPL-bNA (Bachem Inc., 1 mg dissolved in DMSO), 0.1 ml Fast Garnet GBC (Sigma) dye (10 mg/ml in water), 6 ml of 50 mM phosphate buffer, pH 7.0 and 0.2 ml DMSO. The plates were incubated for twenty minutes at room temperature and washed three

times with saline solution (Atlan, *et al.*, 1989; Alvarez, *et al.*, 1985).

Screening of 2,700 colonies using GPL-bNA revealed two colonies which did not turn red. Testing supernatants from liquid cultures of one of these colonies (12-5 or 12-8), with various chromogenic tripeptide substrates (Table V), confirmed that this specific hydrolytic ability had been either eliminated or at least very substantially reduced compared to the original untreated *S.lividans* strain.

Table V

Tripeptidyl Aminopeptidase Activity (A_{540} above background)

Substrate / Supernatants	GPL-bNA	GPM-bNA	APF-bNA	D-FPR-bNA
12-5/T2	0.01	0.01	0	0
12-5/T4	0.10	0.10	0.05	0.06
12-8/T2	0.02	0.02	0	0
12-8/T4	0.13	0.12	0.12	0.08
1-5/T2	0.01	0.01	0.01	0.02
1-5/T4	2.55	Max	Max	0.09

("Max" indicates a A_{540} reading of >3.0)

In a similar experiment to that described above, a L-bNA substrate was used, resulting in the isolation of one mutant (tap$^-$) strain (1-5) from 1500 colonies screened. By comparison, the Tap activity of this mutant strain was unchanged from that of wild type *S.lividans* 66.

Aliquots of each culture supernatant were added to reactions containing 2.5 mg M-CSF GM-CSF and incubated

at 32°C for 2 minutes. The proteins were separated by SDS-PAGE and visualized by Western blotting, using an antiserum raised against the amino terminal 35 amino acids of GM-CSF. At 40 h (T3), the cultures from the *tap* mutants, #11 and #12 contained less activity for converting M-CSF GM-CSF to M-CSF GM-CSF(-3) than those from the *S.lividans*, MS2 and the *tap* mutant, #1.

Protoplasts were prepared from the various *S.lividans* 66 mutants, and were transformed using the M-CSF GM-CSF expression vector pAPO.M-CSF GM-CSF (as described in Canadian Patent Number 1,295,567 and United States Patent Number 5,200,327). The transformed cells were grown in liquid culture and the supernatant fractions were collected following the teaching of Example 11. Aliquots of each culture supernatant were analyzed by SDS-PAGE. The transformants with the *tap* mutants, 12-5 and 12-8 generally showed more intact M-CSF GM-CSF at later time points in the culture than the *S.lividans*, MS2. However, the formation of M-CSF GM-CSF(-3) was not completely eliminated with the *tap* mutants.

* * * * *

The following examples relate to proteases, other than Tap, derived from *Streptomyces*, their DNA sequences and amino acid sequences. These proteases degrade certain substrates under certain conditions. Example 22 describes one such protease, which displayed a significant amino acid sequence homology with the *Bacillus subtilis* protease BPN' [using the BLAST program (Altschul, *et al.*, 1990) to screen the protein sequence databases) and was therefore designated Ssp (Subtilisin-like-protein)]. An improved strain of *Streptomyces* in which this protease is impaired, was created. Southern blot hybridization indicated that Ssp is present in many *Streptomyces* species. Three other proteases, the DNA sequences and deduced amino acid sequences for two of them, are described in Examples 23, 25 and the N-terminal

amino acid sequence of the third protease is indicated in Example 24. A fourth protease, a unique metalloendoproteinase cleaving substrates with X-Pro sequences, requiring Pro residue in P2' sites is described in Example 26.

Example 21. Screening of the *S.lividans* Genomic Library Using APA-bNA

Following the teaching of Example 10, the *S.lividans* 66 genomic library was used to transform protoplasts of the MS7 mutant strain. Transformant colonies were screened with the substrate APA-bNA. Among the thirteen thousand colonies screened, sixteen clones were isolated by virtue of the plasmid-encoded phenotype (colonies appeared red against a background of pale colonies). Plasmid DNA was isolated from these colonies and used to transform *E.coli* competent cells from which larger quantities of plasmid DNA were isolated. Restriction mapping and Southern hybridization analysis enabled the division of the sixteen clones into nine distinct groups. *S.lividans* transformants from each group were grown in liquid medium (TSB-MOPS-Glucose) or modified R2YE (-KH_2PO_4, 1/2 YE) and the cell free broth was assayed for the hydrolysis of chromogenic tripeptide substrates and the degradation of M-CSF GM-CSF. The results are summarized in Table VI.

Table VI
***S.lividans* Clones with APA-bNA Hydrolyzing Activity by Plate Assay**

APA-bNA Activity M-CSF GM-CSF Activity

		APA-bNA Activity			M-CSF GM-CSF Activity	
Group	Examples	Plate	TMP	R2YE*	TMG	RZYE*
1	P5-1 -2 -3 -12	++++	+++++	n. d.	+	n.d.
2	P5-4 -15	++	+++	+++	+	+
3	P5-5	++++	n.d.	n.d.	-	n.d.
4	P5-6 -7 -17 -20	++++	+	+++++ +		+
5	P5-9	++	+	++		-
6	P5-10	++	++	++++	-	+
7	P5-13	+	++	+		n.d.
8	P5-19	+	+	+		n.d.
9	P5-22	++++	+	+++		-

The four plasmid clones from the first group (P5-1, P5-2, P5-3 and P5-12) are identical to P3-5 and P3-13 containing the *tap* gene which were previously isolated using GPL-bNA. Also the single clone P5-5 was previously identified using GPL-bNA as P3-1, P3-3 and P3-6.

The degradation of M-CSF GM-CSF was determined according to the teaching of Example 2. Analysis of *in vitro* degraded M-CSF GM-CSF on native PAGE (FIG. 9) shows no significant M-CSF GM-CSF degradation relative to that of the control strain containing pSS12 (lane 1) with MS7 transformants containing P5-9 (lane 4) from Group 5 and P5-22 (lane 8) from Group 9. However, degradation of M-CSF GM-CSF to its -3 form was observed in modified R2YE cultures of MS7 containing

P5-4 (lane 2) and P5-15 (lane 6) of Group 2, P5-6 (lane 3) and P5-17 (lane 7) of Group 4, and P5-10 (lane 5) of Group 6. Plasmid DNA from each of these groups will be characterized in turn in Examples 22, 23 and 24.

Example 22. Characterization of P5-4 and P5-15

Following the teaching of Example 21, two clones were identified and characterized by restriction mapping as distinct from *tap* and others that were isolated using the APA-bNA substrate. Restriction enzyme site mapping established that two clones (designated P5-4 and P5-15) were shown to represent overlapping fragments of *S.lividans* chromosomal DNA containing the Ssp-encoding gene.

FIGURE 10 shows the restriction enzyme sites present in the P5-4 and P5-15 DNA. The hydrolytic capabilities of strains containing the cloned DNA (or deletions thereof) was measured visually using the agar plate assay method. Southern hybridization against chromosomal DNA showed the expected pattern of hybridizing bands indicating that no major DNA rearrangements had occurred during the isolation of these clones.

Following the teaching of Example 16, the region of DNA encoding the proteolytic activity was defined within the deletion clones P5-4-1 and P5-4-3 (FIG. 10). Specifically, the larger of the two *Nco*1 fragments deleted in P5-4-2, P5-4-4 and P5-4-5 appeared to be correlated with the proteolytic activity.

An SDS-PAGE analysis of protein secreted by strains carrying the P5-4 DNA displayed a major protein band that migrated at a position consistent with a molecular weight of approximately 45,000 daltons. Preparative SDS-PAGE followed by electrotransfer to PVDF membrane (as described in Example 14) allowed direct automated Edman degradation to be carried out to yield the amino acid

sequence NH_2-Asp-Thr-Gly-Ala-Pro^{5}-Gln-Val-Leu-Gly-Gly-^{10}-Glu-Asp-Leu-Ala-Ala-^{15}-Ala-Lys-Ala-Ala-Ser^{20}-Ala-Lys-Ala-Glu-Gly^{25}-Gln-Asp-Pro-Leu-Glu^{30}.

DNA sequence analysis (shown in FIG. 11) of the P5-4 DNA revealed a potential protein coding region located within the region of DNA defined by the two *Nco*I fragments in FIG. 10. This was consistent with the respective activities of the plasmid deletion clones P5-4-1, P5-4-2, P5-4-3, P5-4-4 and P5-4-5. Inspection of the predicted protein sequence revealed the exactly matching, experimentally determined amino terminal amino acid sequence noted above. Furthermore, the predicted amino acid sequence also showed a putative signal peptide at the amino terminus, followed by a putative propeptide defined by the experimentally determined amino terminal sequence of the secreted protease. The amino acid sequences of the proteins predicted from the P5-4 DNA sequence were compared with that of the *Bacillus* protein subtilisin BPN' (FIG. 12), where 1 designates the *S.lividans* sequence while 2 designates the *Bacillus* sequence.

Fermentation yields of Ssp from the MS7 strain carrying the P5-4 plasmid were greater when the medium was supplemented with 5mM Ca^{2+}. The enzyme has been partially purified from a Ca^{2+} supplemented fermentation. After removing the cells by centrifugation, the enzyme precipitated in the 30-70% saturation range with ammonium sulfate producing a fragile, floating pellet. The nature of the pellet made it necessary to decant the supernatant through gauze to get a good yield during collection. Ssp was then subjected to ion exchange chromatography. The enzyme bound to Q-Sepharose at pH 6.0, but the rate of binding was slow. After binding the pH was lowered to 4.0 and the NaCl concentration was between 250 and 500mM for elution. Further purification was obtained by size exclusion chromatography on Superdex 200 at pH 8.0 in

the presence of $CaCl_2$. This step indicated a solution molecular weight of 40-45kd and a band at 43kd was visible in the active fractions after silver staining a reducing Laemmli gel (FIG.2B). This process yielded enzyme which was 80-90% pure. This preparation was used for the characterization below.

The enzyme cleaved APA-pNA, APM-pNA, APA-AMC, and APS-bNA at 800μM substrate. Unlike Tap, Ssp showed no preference for APA-pNA relative to APM-pNA. The proteinase exhibited no activity towards Boc-APS-bNA, N-Ac-APA-pNA, Boc-FSR-AMC, D-PFR-pNA, T-bNA, S-bNA, R-pNA, L-pNA, P-pNA, N-Bz-R-pNA, Boc-AAPA-pNA, N-Ac-AAPA-bNA, N-Suc-AAPF-pNA, Boc-LSTR-pNA, or MAAPV-pNA.

Ssp converted M-CSF GM-CSF to M-CSF GM-CSF(-3). Ssp sequentially removed tripeptide units from the N-terminal of IL-3 generating IL-3(-3), IL-3(-6), IL-3(-9), and IL-3(-12) from intact IL-3. The presence of a disulfide bond at Cys16 may block further processing.

Given the vulnerability of both M-CSF GM-CSF and IL-3 to cleavage by Ssp, the inhibition or elimination of the activity can have a significant impact on product yield and homogeneity from *Streptomyces* fermentations. Ssp is more resistant to APACMK than Tap requiring 81μM APACMK for complete inhibition while Tap was fully blocked by 2μM APACMK. Ssp was also inhibited by PMSF (5mM, complete) and D-FFR-CMK (1mM, partial). The chloromethyl ketones L-CK, F-CK, and TLCK had no effect at 1mM. Unlike Tap, the chelators EGTA and EDTA were effective giving >95% inhibition against enriched Ssp and Ssp in fermentation broth when the concentration of chelator was 0.25mM above the calcium ion concentration. The chelator, Iminodiacetic Acid, was not inhibitory. The following compounds were not inhibitory: captopril (1mM), thiorphan (1mM), iodoacetic acid (1mM), DTT (2.5mM), phosphoramidon (1mM), elastatinal (0.5mM), and pepstatin (0.5mM).

Several conclusions were drawn from the accumulated data. First, Ssp appeared to be a tripeptidyl aminopeptidase which, like Tap, selectively removed tripeptides from substrates possessing a free alpha-amino group. The inhibitor profile was that of a subtilisin type enzyme as suggested by the DNA homology. This enzyme type was characterized by sensitivity to serine proteinase inhibitors and chelating agents. Unlike Tap, Ssp required the divalent cation, Ca^{2+}, for activity. While the use of simple chelating agents during fermentation may seem attractive, we found that many chelators (including EDTA and EGTA) were lethal to *Streptomyces* ; therefore, the development of specific inhibitors or deletion mutants would be of critical importance in the control of this activity during the fermentation.

Deletion of the Ssp-encoding DNA from the *S.lividans* chromosome was accomplished following the teaching of Example 17. Specifically, the DNA from plasmid deletion clone P5-4-4 (FIG. 10) was subcloned into pT7T3 using the *Eco*RI site immediately adjacent to the leftward side of the DNA insert (shown in FIG. 10). Since there was no convenient restriction enzyme site to the rightward side of the DNA insert this was excised using the *Xho*I site (in the replication origin of the plasmid vector, pSS12) which was subsequently ligated to the *Sal*I-digested pT7T3. Hence, overall the *Eco*RI-*Xho*I fragment was inserted in *Eco*RI and *Sal*I digested T7T3 DNA. The fragment was subsequently excised by digestion with *Eco*RI and *Hind*III and inserted into the integration vector, pINT using the same restriction enzyme sites. The pT7T3 intermediate step was required because the *Sal*I site in the multiple cloning site of pINT was not unique and, therefore, not convenient for subcloning purposes.

This integration clone was used to create strains containing the specific deletion at the *ssp* locus in two

S.lividans host strains. Firstly, the MS7 host strain was used to create a new strain designated MS11 (*pepP1-, pepP2-, slpA-, slpC-, tap-, ssp-*). Secondly, another *tap*-deleted strain (MS9) was used to create MS12 (*tap-, ssp-*). The deletion strains MS7, 9, 11 and 12 were cultured in TSB/PPG liquid medium for 22 hours and examined for the ability of cell-free broth to hydrolyse APA-pNA.

Table VII

APA-pNA Hydrolytic Activity of Various *S.lividans* Strains

Strain	ΔA_{405} ml^{-1} min^{-1}
wild type	30.4
MS7	4.5
MS9	2.4
MS11	1.7
MS12	1.3

The results (Table VII) show a reduction in hydrolytic capability, with the MS12 strain showing the lowest activity. All the strains displayed a significantly reduced hydrolytic capability compared to *S.lividans* 66 but the MS9 strain showed a lower level than the MS7 strain. (This was shown in a separate experiment not to be due to the different integration clones used, since MS8 used the same integration clone as MS7 but was derived from *S.lividans* 66 protoplasts and showed indistinguishable properties to MS9.)

Southern hybridization experiments detected DNA sequences homologous to the *ssp* DNA in many *Streptomyces* species. According to the teaching of Example 15, Southern blot hybridization was performed using the 2.25

kb *Bam*HI - *Kpn*I DNA fragment which had been subcloned into pT7T3.18u as a probe. chromosomal DNA was isolated from *S.alboniger*, *S.ambofaciens*, *S.coelicolor*, *S.fradiae*, *S.griseus*, *S.lividans* 66, *S.parvulus*, *S.rimosus* and was digested with either NcoI or SphI.

It should be noted that the same library of clones was screened as in Example 10. Presumably, the lower background level of APA-bNA-hydrolysing activity in MS7 (compared to *S.lividans*) allowed the P5-4 and P5-15 clones to be identified. This has been noticed by other workers particularly relating to neutral protease activities in *B.subtilis* (Sloma, *et al.*, 1990).

Example 23. Characterization of P5-6 and P5-17

Following the teaching of Example 21, four clones were identified and characterized by restriction enzyme mapping as distinct from *tap*, *ssp* or others that were isolated using the APA-bNA substrate. Clone numbers P5-6 and P5-17 were shown to represent overlapping fragments of chromosomal DNA (FIG. 13), wherein P5-7 and P5-20 were identical sibling clones of P5-6.

The common restriction enzyme site map of the P5-6 and P5-17 DNA and deletion clones derived from P5-17 were determined (FIG. 13). Activity against APA-bNA was shown by the number of pluses adjacent to each plasmid and was estimated using the agar plate assay method described in Example 10.

Although these clones encoded significant hydrolytic capability against the APA-bNA substrate in the agar plate assay, no activity above background was observed in cell free broth derived from cultures containing these plasmids grown in TSB media. Neither was it possible to experimentally identify the protein product of this locus. When cultured in liquid medium resembling the agar medium composition (i.e. R2 without added phosphate or agar and containing 0.25% yeast

extract - instead of the usual 0.5%), degrading activity was observed in the cell free broth. However, in contrast to the Tap and Ssp proteins, there was no hydrolysis of GPL-bNA in R2, although there was degradation of full-length M-CSF GM-CSF according to the methods described in Example 2.

DNA sequence analysis of the P5-6 DNA (FIG. 14) revealed a potential coding region. The predicted protein once again displayed a putative secretion signal peptide, followed by a predicted protein of 492 amino acid residues (FIG. 14). Furthermore, when the amino acid sequence was compared to that of the Tap (FIG. 15) a strong homology was obvious around the region encoding the putative active site serine residue. The amino acid sequence for Tap is in the upper row and the amino acid sequence for P5-6 is in the lower row.

Example 24. Characterization of P5-10.

Another cloned DNA fragment was isolated from the same APA-bNA screening experiment described in Example 21. This DNA species was designated P5-10 and showed a different pattern of characteristic restriction enzyme sites (FIG. 16) than those observed for the other clones described above. The protease gene was localized to a 2.25-kb *Bam*HI fragment by deletion subcloning followed by agar plate activity assay against APA-bNA (FIG. 16).

A significant protein band was observed by SDS-PAGE analysis of supernatants of strains carrying this plasmid. Its molecular weight was approximately 50,000 daltons. Amino terminal amino acid sequence analysis was carried out according to the teaching of Example 13 yielding the following sequence : Ala-Glu-Pro-Lys-Ala^{5}-Val-Asp-Ile-Xaa-Asp^{10}-Arg-Leu-Leu-Ser. The activity of supernatant material containing this protein from MS7 host cultures, grown in TSB medium, was very low against APA-bNA and GPL-bNA. However, when cultured in R2YE

liquid medium, a high level of activity was observed against APA-bNA but not GPL-bNA. Furthermore, degradation of full-length M-CSF GM-CSF, according to the methods described in Example 2, was also detectable in samples grown in R2YE but not TSB.

DNA sequence analysis of the 1.9-kb BamHI-SphI fragment of the P5-10 (FIG. 17) revealed a potential protein coding region of 474 amino acids which is consistent with a predicted protein of 53.2 kd. At the amino terminus of the predicted protein sequence there is a 26 amino acid sequence having the properties expected of a conventional signal peptide (van Heijne 1986). The predicted residues 27-40 matched exactly the experimentally determined sequence of the secreted protein. No strong homologies were observed when the predicted protein sequence was compared to protein database sequences, however, weak similarities were noticed to members of the serine carboxypeptidases and prolylendopeptidase families.

Example 25. Characterization of P8-1 and P8-2.

A chromogenic substrate was designed to model the amino terminal region of M-CSF GM-CSF except that the amino terminal residue was modified by the addition of a Boc-group (or other similar moieties such as Fmoc), such that proteases whose activity required a free NH_2-group would be unable to act directly on this substrate. However, any endoprotease present in the *S.lividans* host having a recognition sequence compatible with that of the substrate (specifically Boc-APARSPA-bNA) would be able to cleave and remove the Boc-group in addition to some portion of the peptide. Such cleavage would generate a smaller peptide-linked bNA moiety which would contain a free NH_2-group at the N-terminus and could be acted upon to release the chromogenic bNA moiety which could

subsequently be visualized by reaction with Fast Garnet GBC dye.

This strategy was used to screen the *S.lividans* 66 genomic DNA library after transformation into the MS5 host strain (*tap*+). After screening of eight thousand colonies, six clones were confirmed to encode the ability to degrade the substrate significantly faster than the host strain alone. Two clones proved on restriction enzyme site analysis to be identical to P5-6 described in Example 23. Another clone was similarly shown to be the same as P5-17. Three other clones (P8-1, 2 and 3) were isolated and shown to represent the same region of chromosomal DNA (by Southern hybridization experiments). P8-3 contained a larger DNA fragment which was probably derived from the co-cloning of non-contiguous *Sau*3AI fragments in the construction of the library. P8-1 contained an inserted DNA fragment of approximately 8 Kb, while P8-2 had a smaller insert (3.6 Kb). Deletion mapping followed by agar plate activity assay against Boc-APARSPA-bNA localized the protease gene to a 2.2-kb KpnI-*Bam*HI fragment of P8-1 (FIG. 18). DNA sequence analysis revealed a potential protein coding region of 515 amino acids in the central part of the P8-2 clone (FIG. 19). Comparison of the predicted protein sequence derived from the DNA sequence (FIG. 19) with those encoded by the *tap* and P5-6 clones showed a significant homology between the proteins encoded by P8-2 and P5-6. A smaller but still significant homology was detectable with the Tap protein. Specifically of interest was the conservation of amino acid sequences around the putative active site serine residues of these proteins as follows:

Tap	-	G V S Y G T Y L G A V Y G T L F P D H V R R
P5-6	-	G A S Y G T F L G A T Y A G L F P D R T G R
P8-2	-	G I S Y G T E L G G V Y A H L F P E H V G R

Deletion of P8-1 DNA sequences from the *S.lividans* chromosome was performed according to the teaching of Example 16. Specifically, the DNA from clone P8-1 was digested with *Stu*I and ligated together to form the deletion subclone P8-1-1 which was missing an internal 1.55-kb *Stu*I fragment that encoded the carboxy-terminal half of the protease. The 2.45-kb *Mlu*I-*Bgl*II fragment of P8-1-1 was then ligated into the *Mlu*I and *Bam*HI sites of pINT. The resulting integration vector pINT.P8-1-2 contained the *Stu*I deletion flanked by P8-1 DNA sequence of 1.1 and 1.35 kb. Alternatively, the 0.95-kb *Bam*HI-*Kpn*I fragment of P8-2 and the 1.25-kb *Kpn*I-*Pst*I fragment of P8-1 was ligated into the *Bam*HI and *Pst*I sites of pINT. The resulting integration vector pINT.P8-1-3 has a 3.5-kb *Kpn*I deletion fragment that encoded the entire protease.

These integration vectors were used to create strains containing the specific deletions at the *slpE* locus in *various S.lividans* host strains, by transformation, resolution and hybridization screening following the teaching of Example 16. The strain MS12 was transformed with pINT.P8-1-3 and pINT.P8-1-2 to create the new strains designated MS16 and MS18, respectively (tap$^-$, ssp$^-$, slpE$^-$). The strain MS15 (tap$^-$, ssp, pepN$^-$), which was created from MS12 and an integration vector that was constructed from a DNA clone encoding the *S.lividans* aminopeptidase N gene (*pepN*) (Butler, *et al.*, 1994), was transformed with pINT.P8-1-3 and pINT.P8-1-2 to create

the new strains designated MS17 and MS19 (tap⁻, ssp⁻, pepN⁻, *slpE*⁻). Wild type *S.lividans* 66 was also transformed with pINT.P8-1-2 to create the new strain designated MS20 (*slpE*⁻).

Example 26. X-Pro Metalloendoproteinase Isolation and Characterization

During the production of M-CSF GM-CSF using CANGENUS™, degraded forms of M-CSF GM-CSF, namely M-CSF GM-CSF(-4) and GM-CSF(-6) were identified (FIG. 1A, Lane 4). Significant degradation of IL-3 was also observed (FIG. 1B).

FIGURE 1(B) shows an SDS urea gel electrophoresis (6M urea in the polyacrylamide gel) of IL-3 degradation. A 20-fold concentrated fermentation broth was prepared by subjecting a cell-free fermentation broth to ultrafiltration employing a membrane with a 10 kd cutoff. Lane 2 shows IL-3 before incubation. Lane 1 shows IL-3 after incubation at 32° C.

Initial studies employed wild-type *S.lividans*. The new protease was followed by assaying for the production of M-CSF GM-CSF(-4). A mutant strain (MS12, Example 20) (in which both the Tap and Ssp enzymes were deleted) was the mutant of choice since the absence of GM-CSF(-3) generating activity simplified and improved the sensitivity of the assay for M-CSF GM-CSF(-4) production. The pH optimum, inhibitor profile, substrate specificity, and chromatographic behaviour of the MS12 and wild-type enzymes were essentially identical.

XP-Mep was purified by a combination of chromotography and electrophoresis. The bacteria were removed from the broth containing the enzyme by centrifugation followed by microfiltration through a 0.45μm membrane (Millipore). To reduce aggregation, the broth was not concentrated. The broth was made 1.8M in ammonium sulfate and loaded onto a column of MacroPrep

Methyl, resin (BIORAD) a hydrophobic interaction column. The enzyme was eluted with 100mM Tris-HCl, pH 8.0.

The active fractions were dialysed against 2mM Tris-HCl, pH 8.0 to reduce the ionic strength. The dialysed protein was further purified by preparative isoelectric focusing (ROTOFOR™, BIORAD) employing a gradient from pH 3-10 and a 4% ampholyte loading. The active fractions were pooled and subjected to a second round of preparative IEF to improve the purity and to concentrate the enzyme. Upon cooling, the protease underwent a quantitative isoelectric precipitation. The precipitated enzyme was collected by centrifugation and redissolved in a 20mM buffer with a pH between 7 and 8. While not as precise as analytical IEF, the preparative runs indicated the enzyme has a pI in the 4.2-4.5 range.

The cleavage specificity of XP-Mep was examined using proteins, peptides, and synthetic substrates. The enzyme was capable of cleaving M-CSF GM-CSF, IL-3, IL-6, and gelatin. The cleavage of M-CSF GM-CSF by XP-Mep was followed by native gel electrophoresis and amino acid sequence determination of the products. The enzyme converted intact M-CSF GM-CSF to the M-CSF GM-CSF(-4) and M-CSF GM-CSF(-6) degradation forms. The protease cleaved the Arg^4-Ser^5 (Arg^4-Ser^5-Pro^6 sequence) and Pro^6-Ser^7 (Pro^6-Ser^7-Pro^8) bonds. Interestingly, when purified M-CSF GM-CSF(-3) was incubated with XP-Mep only GM-CSF(-6) was formed. If the concentration of XP-Mep was great enough, M-CSF GM-CSF was degraded to products of less than 10kd in size.

XP-Mep added to the degradation taking place during fermentation. XP-Mep converted M-CSF GM-CSF(-3) to GM-CSF(-6) during fermentation. Tap converted M-CSF GM-CSF to M-CSF GM-CSF(-3) and then was unable to generate further cleavage. After XP-Mep converted the M-CSF GM-CSF(-3) to M-CSF GM-CSF(-6), Tap was then able to further degrade the M-CSF GM-CSF. This enhanced

degradation proceeded to at least the M-CSF GM-CSF(-12) and, possibly, the M-CSF GM-CSF(-15) level. Tap was also capable of converting GM-CSF(-4) to the M-CSF GM-CSF(-7), M-CSF GM-CSF(-10), and, possibly, M-CSF GM-CSF(-13) degradation products. This synergistic effect means that production losses during fermentation may greatly exceed those estimated by measuring the production of M-CSF GM-CSF(-3) and M-CSF GM-CSF(-4).

XP-Mep degraded IL-3. When monitored by isoelectric focusing, the enzyme caused the gradual disappearance of the IL-3 band. When examined by SDS-PAGE, a set of multiple degradation products was clearly visible. The ability of the protease to cleave the IL-3(1-10) peptide was examined. The enzyme cleaved the peptide at the Thr^{6}-Thr^{7} (Thr^{6}-Thr^{7}-Pro^{8}) bond.

Two cleavage sites were identified when IL-6 was the substrate. The protease cut the Ala^{13}-Ala^{14} (Ala^{13}-Ala^{14}-Pro^{15}) and Arg^{17}-Gln^{18} (Arg^{17}-Gln^{18}-Pro^{19}) bonds.

The data obtained with protein substrates suggested that the enzyme cleaved at the N-terminal side of X-Pro sequences.

Several casein derived synthetic peptides were digested with XP-Mep. Beta-casomorphin (YPFPGPI, beta-casein 60-66) was cleaved to yield YPFP and GPI. With VAPFPQV (alpha-casein 25-31) as substrate, FPQV was formed. The peptide PFLQPE (beta-casomorphin 86-91) was cleaved to PFL and QPE. VVPPFLQPE (beta-casomorphin 83-91) was cleaved to VVPPFL and QPE.

A number of potential substrates based on the GM-CSF(-3) structure were synthesized. RAP-pNA, N-Ac-RAP-pNA, and RAPAPA-pNA were not cleaved by XP-Mep. N-Ac-RAPAPA-pNA was cleaved readily by the enzyme liberating N-Ac-RAP and APA-pNA which can be followed by adding Tap.

XP-Mep was tested against a variety of commercial and in-house substrates with no success. The protease failed

to cleave R-pNA, L-pNA, P-pNA, S-bNA, D-PFR-pNA, N-Bz-R-pNA, GR-pNA, pEGR-pNA, N-Bz-VLK-pNA, APA-pNA, SPA-bNA, Boc-APS-bNA, N-(3-carboxypropionyl)-APS-bNA, APPS-bNA, N-Ac-APPT-bNA, Boc-LSTR-pNA, N-Bz-GSHLV-4MbNA, and Boc-FSR-AMC.

The enzyme exhibited the inhibitor profile of a metalloproteinase. The enzyme was strongly inhibited by EDTA. Serine, aspartyl, and cysteinyl protease inhibitors had no measurable effect on activity in the M-CSF GM-CSF assay. The enzyme exhibited some sensitivity to divalent cations. Several more specialized metalloproteinase inhibitors, such as captopril and phosphoramidon, were tested but had no measurable effect on the enzyme activity.

Several other chelating agents were tested for inhibitory effects. Only EDTA and 1.10-phenanthouroline were found to be strongly inhibitory at 1 mM. They were also the most potent inhibitors at pH values at or below neutrality. Partial inhibition was obtained with 10mM 2,3-dimercapto-1-propanesulfonate, 2,3-dimercapto-propanol, meso-2,3-dimercaptosuccinate, triethyl-enetetramine, or EGTA at pH 8.0. No effect was seen with 2',2-Dipyridylamine, 2',2-dipyridyl, salicylic acid, di-2-pyridyl ketone oxime, and iminodiacetic acid.

Divalent cations were tested for inhibitory/stimulatory effects. The protease was stimulated by 0.1-1.0mM Co^{2+}. Higher concentrations of Co^{2+} were inhibitory. Ca^{2+}, Mg^{2+}, and Mn^{2+} had no effect at all. Zn^{2+} was inhibitory above 1mM. Cu^{2+} and Ni^{2+} exhibited inhibitory effects above 5μM.

XP-Mep was completely inactivated by heating at 55°C for 3 hours. The enzyme appears to be stable for 4 days at 37°C. The effect of pH on -4 activity towards M-CSF GM-CSF has been examined. The enzyme was active against M-CSF GM-CSF from pH 5.5 - 8.0 with maximal activity

from 6.0 - 7.0 based upon the recovery of intact M-CSF GM-CSF at the end of the assay. This profile was similar to that obtained with N-Ac-RAPAPA-pNA which exhibited a maximum at 6.0-6.5 and sharp drop in activity below pH 6.0.

The enzyme had a very strong tendency to aggregate. The aggregates were very high in molecular weight and could be readily separated from non-aggregated proteins by size exclusion chromatography. Numerous attempts to break down the aggregate were unsuccessful. While proteolytic activity was stable for at least 2 hours in 4M urea, 0.25M guanidine-HCl, 0.1% SDS, 1% NP40, 1% Tween80, 10mM CHAPS, 2% TritonX100, 1% sodium deoxycholate, 1% sodium taurodeoxycholate, 1.0M NaCl, 25mM beta-mercaptoethanol, or 1mM dithiothoureitol, so was the aggregate. The aggregate and activity were also stable to chloroform extraction, methanol precipitation, lysozyme, trypsin, elastase, DNase, RNase, and lipase treatments designed to eliminate bits of cellular debris that may act as nucleation sites for aggregation.

The non-aggregated protease had an apparent molecular weight of 55-60kd when examined by size exclusion chromatography. The protease gives an apparent molecular weight of 57-60kd when examined by reducing and non-reducing SDS-PAGE.

N-terminal sequencing by Edman degradation after non-reducing SDS-PAGE was conducted. Zymogram studies employing immobilized M-CSF GM-CSF and gelatin were conducted to try to confirm identity of the protease. The proteolytic activity of the band was completely abolished by treatment with EDTA before or after electrophoresis. The N-terminal sequence obtained was X-Ala-Gly-Ala-Pro-Ala-Thr-Glu-Ala-Lys-Leu-Asp-Phe-Ala-Val (in position X, it was not possible to make an unambiguous assignment). The amino acid composition was also determined (Cys and

Trp were not determined) after hydrolysis with 6N HCl (Table VIII).

Table VIII. Amino Acid Composition

Amino Acid	Mol Percentage
Asp + Asn	14.7
Glu + Gln	7.8
Ser	5.3
Gly	14.7
His	1.0
Arg	4.6
Thr	8.5
Ala	12.6
Pro	4.9
Tyr	2.6
Val	8.3
Met	1.1
Ile	2.0
Leu	7.1
Phe	2.2
Lys	2.7

Example 27. Use of Peptide Leaders to Improve Secretion of Exogenous Proteins.

The *Streptomyces* expression vector (APO.H) containing the *aph* promoter followed by the protease B signal peptide (Garvin and Malek, U.S. Patent No. 5,200,327) was used for the construction of expression vectors with the propeptides. It contained unique *NsiI* and Hind III cloning sites for insertion of genes encoding exogenous proteins. The *NsiI* site contains a GCA codon for alanine at the -1 position of the protease B signal peptide. Two tripeptide leaders that were used were Ala-Pro-Ala (designated AP3) and Ala-Pro-Ala-Ala-Pro-Ala which was designated AP6. Oligonucleotides were designed to encode these amino acids and to create a Pst I site which was then used to introduce DNA fragments encoding proteins to be secreted. The pairs of oligonucleotides when annealed formed sticky ends complementary to those of *NsiI* and *Hind* III. The oligonucleotides APA.1 (GCGCCTGCAGCCTA) and APA.2 (AGCTTAGGCTGCAGGCGCTGCA) were used to make the pAP3.H vector by direct ligation to the *Nsil-Hind*III vector

fragment of pAPO.H, containing the *aph* promoter and encoding the protease B signal peptide. Similarly, APA2.1 (GCGCCGGCGGCGCCTGCAGCCTA) and APA2.2 (AGCTTAGGCTGCAGGCGCCGCCGGCGCTGCA) were used to make the pAP6.H vector. Similar to pAP0.H, the two expression vectors pAP3.H and pAP6.H each have a unique PstI site that contains a GCA codon for the carboxy-terminal alanine of each propeptide.

Synthetic DNA sequences were designed by reverse translation of amino acid sequence for stem cell factor (SCF) (Martin, *et al.*, 1990), interleukin-7 (IL-7) (Goodwin, *et al.*, 1989) and erythropoietin (EPO) (Jacobs, *et al.*, 1985) using a codon selection optimized for *Streptomyces*. These DNA sequences and their reverse complements were used for the synthesis of 15 or 16 oligonucleotides which were annealed and ligated together as described in U.S. Patent No. 5,200,327. The completed synthetic genes were then ligated into the *Pst*I and *Hind*III sites of pT7T319U and used to transform *E.coli*. After screening the transformants by restriction analysis of the plasmid DNA, the synthetic genes were determined to be authentic by DNA sequence analysis. The synthetic DNA sequences encoding SCF, IL-7 and EPO are presented in FIGS. 20, 21 and 22 and respectively. The amino acid sequence translated from nucleotides 5 to 496, 5 to 460 or 5 to 502 in the respective DNA sequences are disclosed. The *Pst*I site contained a GCA codon for the alanine at the -1 position, which was compatible with the pAP0.H, pAP3.H and pAP6.H expression vectors.

*Pst*I-*Hind*III DNA fragments encoding SCF, IL-7 and EPO were each ligated to the *Pst*1-*Hind*III vector fragments of pAP0.H, pAP3.H and pAP6.H. DNA from each of the resulting plasmids was used to transform protoplasts of *S.lividans* 66. Single transformant colonies were grown in 15 ml LB (containing 5 μg/ml thiostrepton) seed medium for 3 days. After homogenization the cultures were

inoculated into 1 liter flasks containing 200 ml TSB. Aliquots were removed after 18, 24 and 30 hours of growth at 30°C. The proteins secreted into the culture supernatant fractions (15 μl aliquots) were analyzed by SDS PAGE and visualized by silver staining. The results for the SCF experiments show (FIG. 24) significantly greater protein secretion of SCF by pAP3.SCF and pAP6.SCF than by pAP0.SCF and pAPz.SCF. (The latter encodes a protease B with two amino acid alterations.) The inclusion of the propeptide increased the secretion of SCF approximately 20 fold, IL-7 approximately 10 fold and EPO approximately 5 fold relative to control vectors lacking the propeptides. Each protein was initially secreted with an amino terminal tripeptide or hexapeptide leader. At a later time in the same culture this initial form of each protein was processed to the mature form with the correct amino terminus by the action of the Tap which was secreted into the medium. The amino terminal structure of each of the proteins prevented the Tap from removing any tripeptides from the amino terminus of each mature protein. This invention is applicable to proteins having an amino terminal structure which would prevent Tap digestion of the mature protein. This invention is also applicable to proteins having inefficient signal peptide processing.

Example 28 Use of Proteases to Improve Secretion of Exogenous Proteins

The most common mechanism for the secretion of proteins across biological membranes involves the proteolytic removal of an amino terminal signal peptide with a signal peptidase. Certain amino acids at or near the amino terminus of the mature protein may block or greatly reduce the efficiency of the signal peptidase, leading to lower secretion of the protein. Some proteins are secreted at low levels using CANGENUS™ expression vector APO.H (see Canadian Patent No. 1,295,563,

1,295,566, and 1,295,567, and United States Patent Number 5,200,327.) Some of these proteins contain structural constraints located very close to the amino terminus of the mature protein, such as cysteine residues which are involved in a disulfide bond. This may cause steric hindrance to the signal peptidase, thereby preventing cleavage and subsequent release of the mature protein. In such a case, the efficiency of signal peptide removal is enhanced by insertion at the signal peptidase processing site of amino acids which provide a more flexible structure between the signal peptide and the amino terminus of the mature protein. The additional amino acids could be removed from the amino terminus of the secreted protein by an aminopeptidase. The action of the aminopeptidase is stopped by the amino acid or protein structure at the amino terminus of mature protein. The aminopeptidase may be present in the culture medium into which the protein is being secreted, or may be subsequently added to the secreted protein during the downstream processing.

The present invention teaches a process for increasing the amount of secreted proteins. Such proteins may have a primary amino terminal sequence which imposes certain physicochemical properties and/or conformational properties. These properties may cause steric hindrance and thereby interfere with the processing of the signal peptide. Structural constraints could also include disulphide bonds.

In illustrative embodiments, suitable proteins are interleukin-7 (IL-7), stem cell factor (SCF) and erythropoietin (EPO), which have disulfide bonds involving the amino terminal, second, fourth and seventh amino acids, respectively. A signal peptide which is suitable for use for the secretion of IL-7, SCF and EPO is the 37 amino acid signal peptide from the *Streptomyces*

griseus protease B precursor (Canadian Patent Nos. 1,295,563; 1,295,566 and 1,295,567).

The present invention further describes the use of short propeptides that are multiples of three amino acids in length which, when placed between the signal peptide and the exogenous protein, can increase the level of secreted protein. A peptide leader of either three (APA) or six (APAAPA) amino acids is placed between the protease B signal peptide and the mature protein.

A signal peptidase and a tripeptidyl aminopeptidase (Tap or Ssp) successively remove the protease B signal peptide and the amino terminal peptide leader respectively from proteins secreted from *Streptomyces lividans*. The action of Tap removes peptides from the propeptide, but does not cleave the exogenous protein, due to an amino-terminal primary sequence or structure, such as a disulfide bond, that prevents further degradation activity.

Tap or Ssp are used to remove a propeptide from the amino terminus of a fusion protein comprising an exogenous protein. In a process for the production of an exogenous protein by the secretion of said fusion protein into the growth medium, Tap or Ssp may be initially present in the growth medium, secreted into the medium during growth, or added after growth to a preparation of said fusion protein.

A further feature of this invention is a *Streptomyces* expression system which has a recombinant DNA sequence encoding a natural *Streptomyces* protease, in addition to linked regulatory sequences which can include a promoter sequence, a transcriptional start sequence, untranslated messenger RNA leader sequence, including a ribosome-binding site sequence, a signal sequence, and a transcription termination sequence. A vector of such an expression system, when transformed into a suitable *Streptomyces* host, has the capacity to direct significant

production of a protease due to the multicopy nature of the vector utilized.

Those skilled in the art can use standard methodology to construct recombinant DNA sequences containing various combinations of the aforementioned components of different protease genes, in combination with transcriptional or translational information from non-protease genes: (1) to increase or decrease the production capacity of a *Streptomyces* host for any of the proteases listed; (2) to express exogenous proteins in a *Streptomyces* host or another suitable eukaryotic or prokaryotic host; or (3) to express endogenous proteins in a *Streptomyces* host or another suitable host.

Example 29. Secretion of Soluble Forms of the Enzymes Encoded by P5-6 and P8-2.

No extracellular hydrolytic activity could be observed in liquid cultures of strains carrying the cloned P8-2 DNA sequence of FIG. 19 even when modified R2 liquid medium was used. Moreover, SDS PAGE analysis with silver staining did not detect extracellular proteins of the anticipated sizes in modified R2 liquid cultures of *S.lividans* MS7 carrying the cloned DNA sequences of FIG. 14 (eg. P5-6) or FIG. 19 (eg. P8-2). Although the strains carrying these cloned DNA sequences clearly exhibited hydrolytic activities against their respective substrates on modified R2 agar plates, significant levels of these activities could not be localized to either the intracellular or extracellular fractions.

Consistent with these observations, the amino termini of the potential coding regions of P5-6 and P8-2, unlike conventional signal peptides, contained sequences which matched well with the signal peptidase II consensus sequence characteristic of lipoproteins. As predicted by von Heijne (1989), the signal peptidase II processing would precede the cysteines in the sequence LATACSAGGAS of P5-6 (FIG. 14) and LTAGCSGGSS of P8-2 (FIG. 19). Each

sequence showed a striking clustering of turn-producing amino acids following the cysteine, consistent with the amino termini of lipoproteins. The highly positively charged amino terminus of the potential coding region of P5-6, with 7 arginines and a single aspartate, has been commonly found on other Gram positive signal peptides. Overall, the amino-terminal sequences for the potential coding regions of P5-6 and P8-2 were consistent with membrane bound forms of each enzyme, designated *Slp*D and *Slp*E, respectively.

In order to allow biochemical purification of the predicted proteins from culture supernates, to examine their hydrolytic capabilities and to confirm that the predicted proteins were directly responsible for these activities, the nucleotides encoding both the putative promoter region and the lipoprotein signal peptide including the +1 cysteine were replaced by sequences for the aminoglycoside phosphotransferase (*aph*) promoter and encoding the protease B signal peptide and the six amino acid propeptide, as described in Example 27. This was accomplished by the use of oligonucleotides to adapt the *Slp*D and *Slp*E proteins at their amino-termini with appropriate cloning sites, for ligation into the expression vector pAP6.H.

To adapt the N-terminus of the *Slp*D protein, oligonucleotides encoding the 11 amino acids of *Slp*D immediately downstream of the SPase II +1 cysteine were synthesized. An *EcoR*I cloning site at the 5' end allowed for ligation of the oligonucleotides into the *EcoR*I site contained within the polylinker of a T7T318U based subclone (#4) of *Slp*D clone p5-6. This subclone also contained a *Hind*III site from the polylinker located 380 nucleotides downstream of the *Slp*D stop codon. The oligonucleotides also contained at their C-terminus a BamH I site, which joined to a natural BamHI site within

the *Slp*D encoding sequence, located 30 nucleotides downstream from the SPase II +1 cysteine.

A subclone containing these oligonucleotides was subjected to DNA sequence analysis, a routine procedure employed to confirm the fidelity of the cloned oligonucleotide sequence, and the sequence was found to be correct. An *Nsi*I cloning site contained within the N-terminus of the oligonucleotides allowed for ligation to the Pst I site of AP6.H and subsequent joining of the protease B signal plus leader directly to the *Slp*D at the serine residue immediately adjacent to the SPase II cysteine. The 1920 *Nsi*I to *Hind*III fragment encoding *Slp*D was subsequently cloned into AP6.H to produce AP6.*Slp*D.

An analogous strategy was used to adapt the N-terminus of the *Slp*E protein with oligonucleotides encoding the 35 amino acids of *Slp*E immediately downstream of the SPase II +1 cysteine. A *Pst*I compatible site located at the 5' end allowed for ligation of the oligonucleotides into the *Pst*I site located within the polylinker of a T7T318U based subclone (#5) of *Slp*E clone p8-2. The oligonucleotides also contained at their 3' end a *Pfl*MI site which joins to a natural *Pfl*MI site within the *Slp*E encoding sequence, located 100 nucleotides downstream from the SPase II +1 cysteine. At the 3' end of one of the oligonucleotides creating the *Pfl*MI site, there was a potential secondary structure which could potentially have caused difficulties in cloning by forming a relatively stable hairpin, thus providing the *Pfl*MI sticky end from participating in the ligation. The sequence of this oligonucleotide and its complement were modified to abolish the hairpin structure, while still encoding the correct amino acid sequence for *Slp*E.

DNA sequence analysis of two of the three pT7T3.18U subclones containing these oligonucleotides showed that their 5' ends did indeed contain the nucleotide sequences

from the oligonucleotides (*i*.e., they contained an *Nsi*I site). Surprisingly, however, the sequences at their 3' ends upstream of the *Pfl*MI cloning site, where the nucleotides should have been substituted to abolish the potential hairpin structure, contained wild type nucleotides. The *Slp*E encoding sequence remained completely intact and in the correct reading frame, and sequences past the *Pfl*MI site also were intact and in the correct frame.

An *Nsi*I cloning site contained within the N-terminus allowed for the subsequent ligation in the correct reading frame into the *Pst*I site of AP6.H and the joining of the protease B signal plus leader directly to the *Slp*E at the serine residue immediately adjacent to the SPase II +1 cysteine. A *Sac*I site located 238 nucleotides downstream of the *Slp*E stop codon was used in conjunction with a *Hind*III - *Sac*I 8mer adapter (AGCTAGCT) to join the 3' end of the *Slp*E clone to the *Hind*III site in the AP6.H expression plasmid. The 1820 bp *Nsi*I to *Sac*I fragment encoding *Slp*E was then used along with the *Hind*III - *Sac*I adapter in a three-way ligation into AP6.H to produce AP6.*Slp*E.

When these plasmids were used to transform protoplasts of MS11, secreted proteins for both AP6.*Slp*D and AP6.*Slp*E were observed at approximate molecular weights of 55kd and 56kd, respectively. Direct automated N-terminal Edman degradation analysis of the secreted proteins produced the following amino acid sequences: SAGGASTXAG for *Slp*D and APAAPASGGSSDEDK for *Slp*E. For *Slp*D, culture supernatants showed a dramatic increase in the ability to hydrolyse APA-bNA.

Table IX
Soluble Protease Substrate Assays

Transformant	Timepoint	A_{405}	A_{540}
SS12	18	0.144	0.100
	23	0.132	0.038
	41	0.126	0.018
p5-6	17.5	1.147	0.246
	23	0.990	0.278
	41	0.105	0.000
p8-2	17.5	0.115	0.084
	23	0.111	0.015
	41	0.108	0.036

The A_{405} values reflected the APA-bNA assay on 20 μl cell free broth from Tap-deleted *S. lividans* 66 cultures. The A_{540} values reflected the Boc-APARSPA-bNA assay on 20 μl cell free broth from *S. lividans* 66 cultures. There was no adjustment for dry weights.

This correlated with the N-terminal sequence data on *Slp*D which showed that it was lacking the leader peptide APAAPA, which may have been cleaved due to autocatalytic activity of the *Slp*D itself. In contrast, *Slp*E culture supernatants showed no ability to hydrolyse APA-bNa, correlating with the presence of an intact P6 leader at the N-terminus of the secreted protein.

Example 30. Use of Tap for the Processing of the Propeptide from Secreted SCF

The Tap protein was purified from a liquid culture of *S.lividans* according to the teaching of Example 1. Alternatively, Tap can be expressed by transformants of *S.lividans* 66 harbouring the plasmid pCAN94 (P3-5) which contains the entire *tap* gene, including promoter/ regulatory sequences, a protein coding sequence, and a putative transcription termination sequence. Tap expression can be further enhanced by the use of a plasmid, pCAN155, in which the *tap* gene, excluding natural sequence upstream of its transcription start

point, was placed downstream of the strong, constitutively expressed *ermE** promoter. pCAN155 was constructed via 3-way ligation of the large *Bam*HI - *Hind*III fragment of pCAN149, and the 0.875 kb *Bam*HI - *Kpn*I and 1.8 kb *Kpn*I - *Hind*III fragments of pPepT, as described. pCAN149 is a derivative of the pAPO.M-CSF GM-CSF expression plasmid in which the small *Bam*HI - *Mlu*I fragment was replaced by a *Bgl*II - *Mlu*I fragment containing the *ermE** promoter and a sequence encoding the amino terminus of the protease signal peptide. pPepT is a derivative of pT7T318U (Pharmacia) into which a 3.0 kb fragment *Pst*I - *Bam*HI from pCAN94 was subcloned. A partial *Bam*HI digest was required to generate this fragment due to the presence of other internal BamHI sites.

pCAN155 and pCAN94 were transformed into *S.lividans* 66 strain MS11 (*tap⁻,ssp⁻*) and grown in liquid TSB medium. Secreted proteins were TCA-precipated from the extracellular broth at 17.5, 23.5 and 41.5 hours after inoculation, and analyzed by SDS PAGE. The level of secreted 55 kd Tap protein was substantially higher from pCAN155 than from pCAN94 (P3-5), indicating the relative promoter strengths of *ermE** and *tap*, respectively. The secreted Tap protein continued accumulate until 41.5 hours of culture to a level of approximately 2 g/l.

Appropriate dilutions of cell-free broth samples collected at various time points from independent cultures harbouring pCAN94, and pCAN155 were tested for ala-pro-ala pNA hydrolysis as described previously. The results are shown in Table X.

Table X

Ala-Pro-Ala-pNa Hydrolysis Activity

Time Point (h)	pCAN155 (ΔA_{405}/min/ml)	pCAN94 (ΔA_{405}/min/ml)
17.5	257	12
23.5	2922	286
41.5	56500	10500
71.5	50000	10000

Table IX illustrates that at all 3 time points the pCAN155 transformants expressed the highest amount of secreted product, followed by the pCAN94 transformants. The pCAN138 transformants represented a control culture of *S.lividans* MSII essentially negative for Tap activity.

Secreted SCF from culture supernatants of the MS12 strain harboring the pAP6.SCF vector were digested with a preparation of Tap to remove the amino-terminal hexapeptide. Aliquots (10μl) of the culture supernatants, containing approximately 100 ng of APAAPA-SCF, were mixed with 0, 1 and 2μl of Tap (200ΔA^{405}/min/ml) and incubated at 37°C for 1 to 3 hours. The proteins were then analyzed by SDS PAGE with silver staining. The results indicated that Tap digestion converted the APAAPA-SCF to APA-SCF and SCF with increasing incubation time and enzyme concentration.

Example 31. An Immunoassay Using Tap.

Tap as a unique protease with a well established assay using a synthetic substrate for determination of its activity (described in this patent application) may be applied as a useful tool for making an immunoassay. In microwell ELISA, antigens were immobilized in a microwell and probed by labelled antibody (conjugate). The enzyme-labelled reagents were detected with the appropriate substrate, which was converted to a visible colored product at the reaction site. The intensity of color produced was proportional to the amount of measured antigen.

To date, the most common enzymes used for generating color have been alkaline phosphatase or horseradish peroxidase. In this example, those enzymes were replaced with Tap and using synthetic substrates, developed and described in this patent application, such as APA-pNA for visible color and APA-AMC for fluorescence technology detection.

To demonstrate this invention, IL-3 was used as an example for antigen quantitation. Rabbit anti-IL-3 antisera (Cangene Corporation) was used as the first antibody. The second antibody, goat anti-rabbit IgG linked to biotin (Sigma), and streptavidin (Boehringer Mannheim GmbH) were used as the amplification system. Tap linked to biotin was used as the enzyme. The Tap was purified as described in Example 1 and 9.0 ml of the Tap (approximately 0.3 mg/mL) were biotinylated with D-Biotinyl-E-aminocaproic acid N-hydroxysuccinimide ester as described in Biochemia Bulletin of Boehringer Mannheim (1989, Antibodies and Reagents for Immunochemistry, at page 115). Serial dilutions of recombinant hIL-3 (Cangene Corporation) were applied to the microplate wells (100 μL/well), and then incubated at 4°C for over 16 hours. The wells then were washed and 5% BSA (bovine serum albumin) was added as a blocker. After

1 hour incubation, the wells were washed and rabbit anti hIL-3 sera (Cangene Corporation) was added at a dilution of 1/2000. Incubation was performed at 37° C for 1 hour. The wells were then washed and the second antibody, goat anti-rabbit IgG-Biotin (Sigma), was added at a dilution of 1/2000 for 1 hour at 37° C. After washing, a mixture of Streptavidin and Biotin-Tap was added. This mixture was prepared previously as follows: 40 μL of Streptavidin (Boehringer Mannheim, 1 mg/mL) and 35 μL of Biotin-Tap were added to 5 mL Tris buffer pH 8.0 containing 1% BSA. The mixture was pre-incubated for 45 minutes before being added to the microplate assay. The mixture was washed from the microplate after incubation for 45 minutes at room temperature. Then 100 μL of the enzyme substrate (0.8 mM) were added. For color development, APA-pNA was used as a substrate and the absorbance was read after 2 and 16 hours incubation at 405nm. For faster analysis, APA-AMC (Cangene Corporation) was used as a fluorescent substrate, where the incubation was performed for 30 minutes and the assay was analyzed at excitation/emission of 400/450nm by the multiwell plate scanning fluorescent system using a Fluorescence Concentration Anaylzer (Pandex).

A hIL-3 calibration curve was generated (FIG. 23) using ELISA technology with Tap as the enzyme and APA-pNA as the substrate for color forming (Panel A) incubated for either 2 hours (o---o) or 16 hours (Δ----Δ), and APA-AMC as a fluorescent substrate (Panel B) incubated for 30 minutes.

* * * * * * * * * *

The present invention has been described in terms of particular embodiments found or proposed by the present inventors to comprise preferred modes for the practice of the invention. It will be appreciated by those of skill in the art that, in light of the present disclosure, numerous modifications and changes can be made in the particular embodiments exemplified without departing from the intended scope of the invention. For example, due to codon redundancy, changes can be made in the underlying DNA sequence without affecting the protein sequence. Moreover, due to biological functional equivalency considerations, changes can be made in protein structure without affecting the biological action in kind or amount. All such modifications are intended to be included within the scope of the appended claims.

CITED DOCUMENTS

The documents listed below are incorporated herein by reference to the extent that they supplement, explain, provide a background for, or teach methodology, techniques and/or compositions employed herein.

Altschul, S.F., W. Gish, W. Miller, E.W. Myers and D. Lipman. 1990. Basic Local Alignment Search Tool, J. Mol. Biol. 215: 403-410.

Ashe, B.M. and Zimmennan (1977) Biochem. Biophys. Res. Commun. 75:194-199.

Atlan, D., P. Laloi and R. Portalier. 1989. Isolation and characterization of aminopeptidase-deficient *Lactobacillus bulgaricus mutants*. Appl. Env. Microbiol. 55:1717-1723.

Alvarez, N.G., C. Bordallo, S. Gascon and P.S. Rendueles. 1985. Purification and characterization of a thermosensitive X-prolyl dipeptidyl aminopeptidase from *S. cerevisiae*. BBA 832:119-125.

Aretz, W., K-P. Koller and G. Riess. 1989. Proteolytic enzymes from recombinant *Streptomyces lividans* TK24. FEMS Microbiol. Lett. 65:31-36.

Balow, R-M., Tomkinson, B., Ragnorsson, U. and Zetterqvist, O. 1986. J. Biol. Chem. Purification, Substrate Specificity and Classification of Tripeptidyl Peptidase II. 261 (5) 2409-2417.

Bender, E., K-P. Koller and J.W. Engels. 1990a. Secretory synthesis of human interleukin-2 by *Streptomyces lividans*. Gene 86:227-232.

Bender, E., Vogel, R., Koller, K.P. and J.W. Engels. 1990b. Synthesis and Secretion of Hirudin by *Streptomyces lividans*. *Appl. Microbiol. Biotechnol.* 34:203-207.

Bibb, M.J., M.J. Bibb, J. M. Ward and S.N. Cohen. 1985. Nucleotide sequences encoding and promoting expression of three antibiotic resistance genes indigenous to *Streptomyces*. *Mol. Gen. Genet. 199:26-36.*

Bibb, M.J., P.R. Findlay and M.W. Johnson. 1984. The relationship between base composition and codon usage in bacterial genes and its use for the

simple and reliable identification of protein-coding sequences. Gene 30:157-166.

Brawner, M., D. Taylor and J. Fornwald. 1990. Expression of the soluble CD-4 receptor in *Streptomyces. J. Cell. Biochem., supplement 14A p103.*

Butler, M.J., C.C. Davey, P. Krygsman, E. Walczyk, and L.T. Malek. 1992. Cloning of genetic loci involved in endoprotease activity in *S.lividans* 66: a novel neutral protease gene with an adjacent divergent putative regulatory gene. Can. J. Microbiol., in the press.

Butler, M.J., J.S. Aphale, M.A. DiZonno, P. Krygsman, E. Walczyk, L.T. Malek. 1994. Intracellular Aminopeptidases in *Streptomyces lividans* 66. J. Indust. Microbiol. 13:24-29.

Davies, B.J. 1964. Ann, N.Y. Acad. Sci. 121, 404

Diedrichs, K., Boone, T., and Karplus, P.A. (1991) Science 254:1779-82.

Doebber, T. W., Divor, A. R. , and Ellis, S. (1978) Endocrinology 103:1794-1804.

Doggette, P.E., and F.R. Blattner. 1986. Personal access of sequence databases on personal computers. Nucleic Acids Res. 14:611-619.

Fink, A.L and Meehan, P. (1979) Proc. Nat'l Acad. Sci. USA 76:1566-69.

Fornwald, J.A., Donovan, J.J., Gerber, R., Keller, J., Taylor, D.P., Arcuri, E.J. and Brawner, M.E. 1993. Soluble forms of the human T cell receptor CD4 are efficiently expressed by *Streptomyces lividans.* Bio/Technology 11:1031-1036.

Fukusawa, K.M. and M. Harada. 1981. Purification and properties of dipeptidyl peptidase IV from *Streptococcus mitis* ATCC 9811. Arch. Biochem. Biophys. 210:230-237.

Hanson, H. and M. Frohne. 1976. Crystalline leucine aminopeptidase from lens in proteolytic enzymes (Ed., L. Lorand) Methods Enzymol. 45:504-521.

Henderson, G., P. Krygsman, C.J. Lui, C.C. Davey and L.T. Malek. 1987. Characterization and structure of genes for proteases A and B from *Streptomyces griseus.* J. Bacterial. 169:3778-3784.

Hopwood, D.A., M.J. Bibb, K.F. Chater, T. Kieser, C.J. Bruton, H.M. Kieser, D.J. Lydiate, C.J. Thompson, C.P. Smith, J.M. Ward and H. Schourempf. 1985. Genetic manipulation of *Streptomyces*, a laboratory manual. The John Innes Foundation, Norwich, U.K.

Hunkapiller, M. W., Forgac, M.D., and Richards, J.H. (1976) Biochem. 15: 5581-85.

Illingworth, C., Larson, G. and Hellekant, G.. 1989. Secretion of the sweet-tasting plant protein thaumatin by *Streptomyces lividans*. *J. of Industrial Microbiology* 4:37-42.

Ingram, C., M. Brawner, P. Youngman and J. Westphaling. 1989. *xylE* functions as an efficient reporter gene in *Streptomyces* spp.: Use for the study of gal P1, a catabolite-controlled promoter. J. Bacteriol. 177:6617-6624.

Koller, K.P., Riess, G., Sauber, K., Uhlmann, E. and Wallmeier, H.. 1989. Recombinant *Streptomyces lividans* secretes a fusion protein of tendamistat and proinsulin. Biotechnology 7:1055-1059.

Kreiger, T.J., Bartfeld, D., Jenish, D.L., Hadary, D. 1994. Purification and characterization of a novel tripeptidyl aminopeptidase from *Streptomyces lividans* 66. FEBS Letters 352: 385-388.

Kreil, G. 1990. Processing of precursors by dipeptidyl aminopeptidases: a case of molecular ticketing. TIBS. 15:23-26.

Lichenstein, H., Brawner, M.F., Miles, L.M., Meyers, C.A., Young, P.R., Simon, P.L. and Eckhardt, T. 1988. Secretion of interleukin-1β and *Escherichia coli* galactokinase by *Streptomyces lividans*. Gene. 129:129-134.

Lloyd, R.J. and G.G. Pritchard. 1991. Characterization of X-prolyl dipeptidyl aminopeptidase from *Lactococcus lactis* subsp. *lactis*. J. Gen. Microbiol. 137:49-55.

Malek, L.T., G. Soostmeyer, C.C. Davey, P. Krygsman, J. Compton, J. Gray, T. Zimny and D. Stewart. 1990. Secretion of Granulocyte Macrophage Colony Stimulating Factor (M-CSF GM-CSF) in *Streptomyces lividans*. J. Cell. Biochem., supplement 14A, p127.

Maniatis, T., E.F. Fritsch and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

McDonald J.K., Hoisington, A.R. and Eisenhauer, D.A. 1985. Partial Purification and Characterization of an Ovarian Tripeptidyl Peptidase: A lysosomal exopeptidase that sequentially releases Collagen-related (Gly-Pro-x) Triplets. 126:63-71.

Menn, F-M., Zylstra, G.J. & Gibson, D.T. 1991. Location and sequence of the *tool F* gene encoding 2-hydroxy -6-oxohepta-2, 4-dienoate hydrolase in *Pseudom... putide* F1. Gene 104:91-94.

Pearson, W.R. and D.J. Lipman. 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA. 85:2444-2448.

Proudfoot, A.E.I, Brown, S. C., Bernard, A.R., Bonnefoy, J. -Y., and Kawashima, E.H. (1993) J. Protein Chem. 12: 489-97.

Schoellmann, G. and Shaw, E. 1963. Direct evidence for the presence of histidine in the active center of chymotrypsin. Biochemistry 2:252.

Shaw, E., Mares-Guia, M., and Cohen, W. 1975. Evidence for an active center histidine in trypsin through the use of a specific reagent, TLCK, the chloromethyl ketone derived from N-tosyl-lysine. Biochemistry 4:2219.

Tagakuchi, S., I. Kumagai, J. Nakayama, A. Suzuki and K. Miura. 1989. Efficient extracelluar expression of a foreign protein in *Streptomyces* using secretory protease inhibitor (SSI) gene fusions. Biotechnology 7:1063-66.

Tinoco, I., Jr., P.N. Borer, B. Dengler, M.D. Levine, O.C. Uhlenbech D.M. Crothers and J. Gralla. 1973. Improved estimation of secondary structure in ribonucleic acid. Nature New Biol. 246:40-41.

Tomkinson, B. and Jonsson, A-K. 1991. Characterization of cDNA for Human Tripeptidyl Peptidase II: The N-Terminal Part of the Enzyme is Similar to Subtilisin. Biochemistry 30:168-174.

Ueda, Y, Tsumoto, K., Watanabe, K. and Kumagai, I. 1993. Synthesis and expression of a DNA encoding the Fv domain of an anti-lysozyme monoclonal

antibody, HyHEL10, in *Streptomyces lividans*. *Gene*. 120:129-134.

von Heijne, G. 1989. The structure of signal peptides from bacterial lipoproteins. *Protein Engineering* 2:531-534.

Wilbur, W.J. and D.J. Lipman. 1983. Rapid Similarity searches of nucleic acid and protein data banks. Proc. Natl. Acad. Sci. USA. 80:726-730.

Yoshimoto, T., N. Murayama, T. Honda, H. Tone, and D. Tsuru. 1988. Cloning and expression of aminopeptidase P gene from *E. coli* HB101 and characterization of expressed enzyme. J. Biochem. 104:93-97.

Yoshimoto, T., H. Tone, T. Honda, K. Osatomi, R. Kobayashi, and D. Tsuru. 1989. Sequencing and high expression of aminopeptidase P gene from *E. coli* HB101. J. Biochem. 105:412-416.

Canadian Patents No. 1,295,563, No. 1,295,566 and No. 1,295,567.

United States Patent No. 5,200,327.

PCT/US92/01598 (Sloma, *et al.*)

We Claim:

1. An isolated and purified, endogenous protease of *Streptomyces*.
2. The protease of claim 1, wherein the protease is a tripeptidyl aminopeptidase.
3. The protease of claim 1, wherein the protease is a metalloendoproteinase.
4. The protease of claim 1, having an amino acid sequence as set out in FIG. 5, 11, 14, 17 or 19.
5. The protease of claim 3, wherein the protease cleaves substrates with X-Pro sequences, where the Pro residue is in the P2' site.
6. An isolated DNA molecule encoding an endogenous protease of *Streptomyces*.
7. The isolated DNA molecule of claim 6 having a nucleic acid sequence according to FIG. 5, 11, 14, 17 or 19.
8. An isolated DNA molecule encoding an impaired, endogenous protease of *Streptomyces*.
9. The isolated DNA molecule of claim 8, wherein said molecule lacks all or part of the sequence between site 2 and site 8 of the nucleotide sequence having the restriction map of FIG. 6.
10. The isolated DNA molecule of claim 8, wherein said molecule lacks all or part of the sequence between site 5 and site 7 of a nucleotide sequence having the restriction map of FIG. 6.
11. A genetic expression system comprising an impaired endogenous protease of *Streptomyces*.
12. The genetic expression system of claim 11, further comprising a recombinant molecule encoding an exogenous gene product and wherein the system is capable of producing the product.
13. A genetic expression system comprising a recombinant DNA molecule which includes

(a) a DNA molecule of claim 6 and 7 and (b) a regulatory sequence operatively linked to the DNA molecule, said system capable of producing the protease encoded by said DNA molecule.

14. A genetic expression system comprising a recombinant DNA molecule which includes all or part of a nucleotide sequence capable of regulating a DNA molecule encoding the protease of claim 1 and wherein said system is capable of producing the protease.

15. A vector including the DNA molecule of claim 6 or 7.

16. A prokaryotic or eukaryotic host cell transformed or transfected with the recombinant DNA molecule of claim 6, 7, 8, 9 or 10.

17. A method for producing an exogenous gene product by a suitable host cell, said method comprising culturing the genetic expression system of claim 12 under suitable conditions and recovering the exogenous gene product.

18. A method for producing a protease by a suitable host cell, said method comprising transforming the host cell with the vector of claim 15 and placing the transformed cell in conditions that allow production of the protease.

19. The method of claim 17 or 18, wherein the host cell is a *Streptomyces* host cell.

20. A protease inhibitor comprising L-alanyl-L-prolyl-L-alanine chloromethylketone.

21. A method of making the inhibitor of claim 20, comprising the steps of:

(a) extending a chain of Boc-AP to obtain Boc-APA-OBz;

(b) de-blocking the C-terminal end of Boc-APA-OBz to obtain Boc-APA; and

(c) simultaneously extending C-terminal end, and de-blocking the N-terminal end, of Boc-APA to obtain APA-CMK.

22. A strain of *Streptomyces* displaying impaired expression of at least one endogenous protease and comprising a DNA molecule capable of encoding the protease.

23. The strain of *Streptomyces* of claim 22, wherein the endogenous protease is tripeptidyl aminopeptidase.

24. The strain of claim 22, selected from the group comprising *S.lividans, S. ambofaciens, S. coelicolor, S. alboniger, S. parvulus* and *S. rimosus*.

25. The strain of claim 22, which expresses an exogenous gene product.

26. The strain of claim 22, wherein a DNA molecule encoding for the protease lacks all or part of the molecule between site 2 and site 8 of the molecule having the restriction map of FIG. 6.

27. The strain of claim 22, wherein the DNA molecule encoding for the protease lacks all or part of the molecule between site 5 and site 7 of the molecule having the restriction map of FIG. 6.

28. The strain of claim 22, wherein the impaired expression decreases the quantity or activity of the endogenous protease.

29. The strain of claim 22, wherein the impaired expression increases the quantity, quality or stability of the exogenous gene product.

30. The strain of claim 22, wherein the protease is impaired by one of the following: deleting nucleotides in the nucleic acid molecule encoding for the protease, deleting and substituting nucleotides in a nucleic acid molecule encoding

for the protease mutating nucleotides in the nucleic acid molecule encoding the protein.

31. A kit for an enzyme-linked immunosorbent assay, said kit comprising in separate containers, the protease of claim 1 covalently linked to a carrier, and a substrate which when cleaved by the protease generates a detectable and measurable signal.

32. A method for improving the expression and secretion of a mature protein from a genetic expression system, said method comprising the steps of:

(a) in the system, expressing a construct that comprises a nucleotide molecule encoding a polypeptide comprised of (i) a first sequence encoding a precursor of the mature protein and (ii) a second sequence encoding a substrate for a protease, wherein said second sequence is added to the N-terminus of said first sequence, such that said polypeptide is obtained; and then

(b) exposing said polypeptide to the protease such that the second sequence is cleaved from the first sequence to form the mature protein, wherein steps (a) and (b) are carried out under conditions for secretion of the mature protein.

33. The method of claim 32, wherein the second sequence is a tripeptide.

34. The method of claim 33, wherein the tripeptide is Ala-Pro-Ala.

35. The method of claim 32, wherein the second sequence is Ala-Pro-Ala-Ala-Pro-Ala.

36. A polypeptide comprising (a) a leader which is a substrate for a signal peptidase, (b) a

propeptide which is a substrate for a protease, and (c) an exogenous protein.

37. The polypeptide of claim 36, wherein the structure of the exogenous protein constrains degradation of the exogenous protein by the protease.

38. The polypeptide of claim 36, wherein, absent the propeptide, the structure of the exogenous protein constrains cleavage of the leader sequence by the signal peptidase.

39. The polypeptide of claim 36, wherein the protease cleaves the propeptide from the polypeptide, thereby yielding an isolated exogenous protein.

40. A use of the protease of claim 1 to remove an amino acid, peptide or polypeptide from a substrate.

41. A use of the protease of claim 3 to digest a protein of connective tissue.

42. A use of the protease of claim 40, wherein the peptide is a peptide leader.

43. A use of the protease of claim 40, wherein the polypeptide is a fusion protein.

FIG. 1

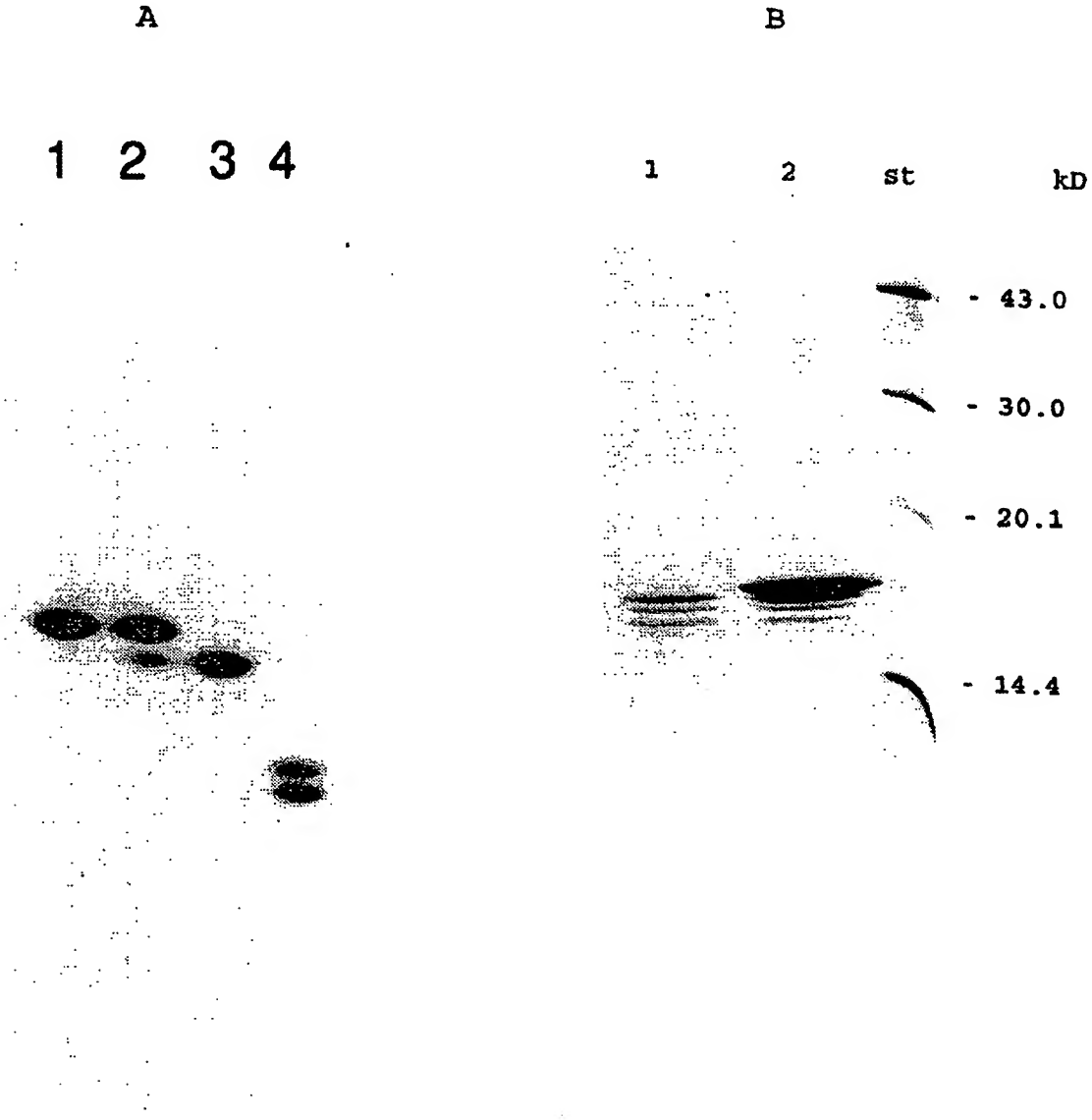


FIG. 2

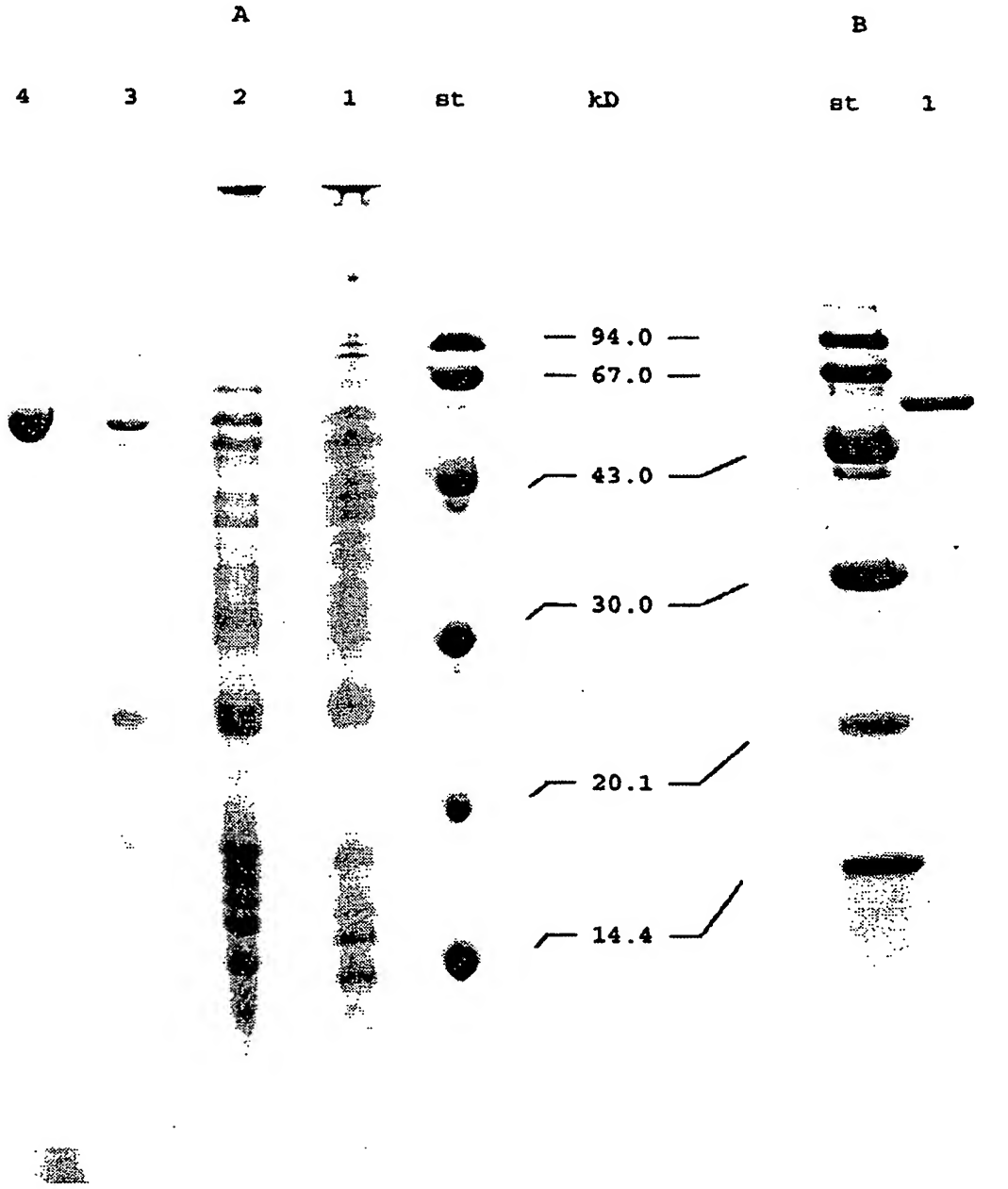


FIG. 3

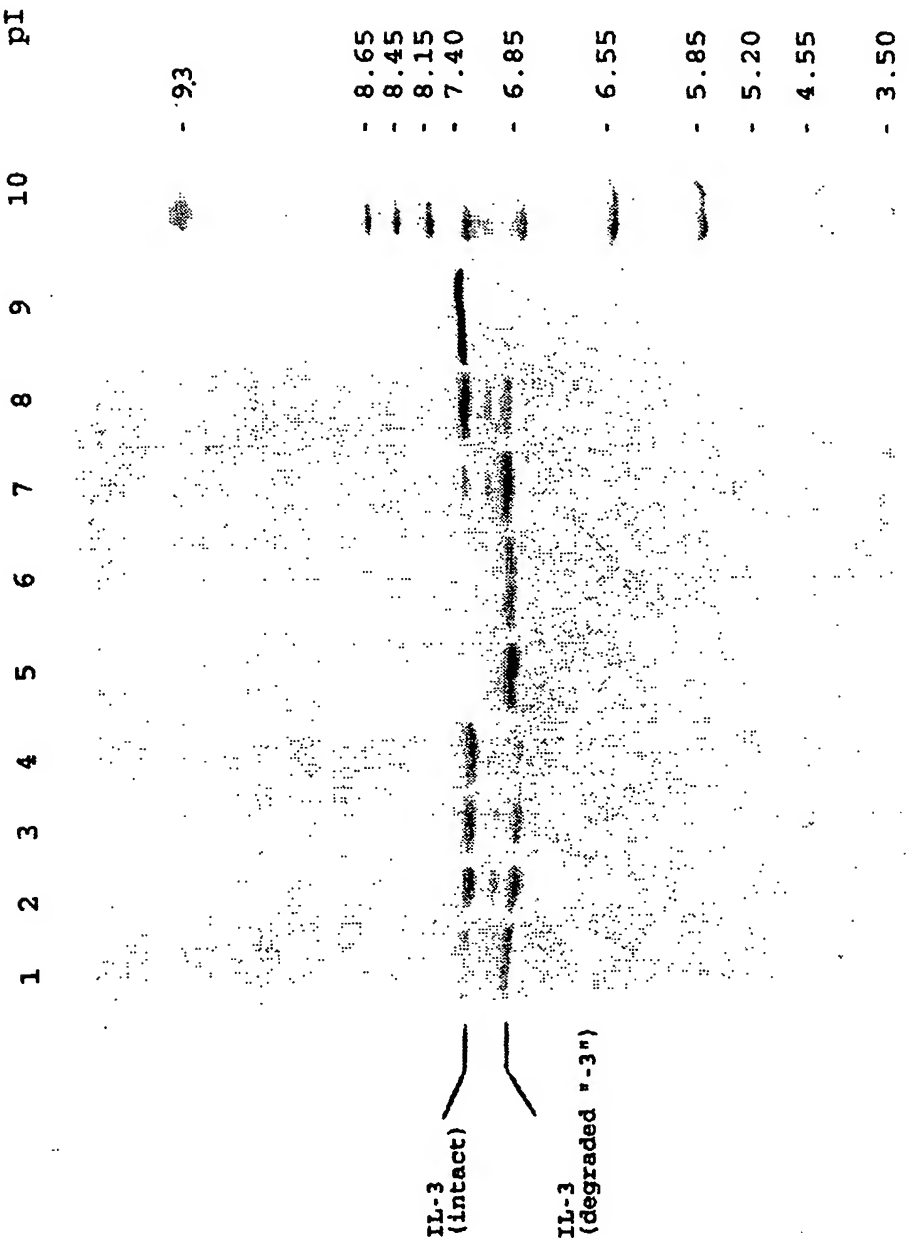

pI
1
2
3
4
5
6
7
8
9
10
9.3
8.65
8.45
8.15
7.40
6.85
6.55
5.85
5.20
4.55
3.50
IL-3 (intact)
IL-3 (degraded "-3")

FIG. 4

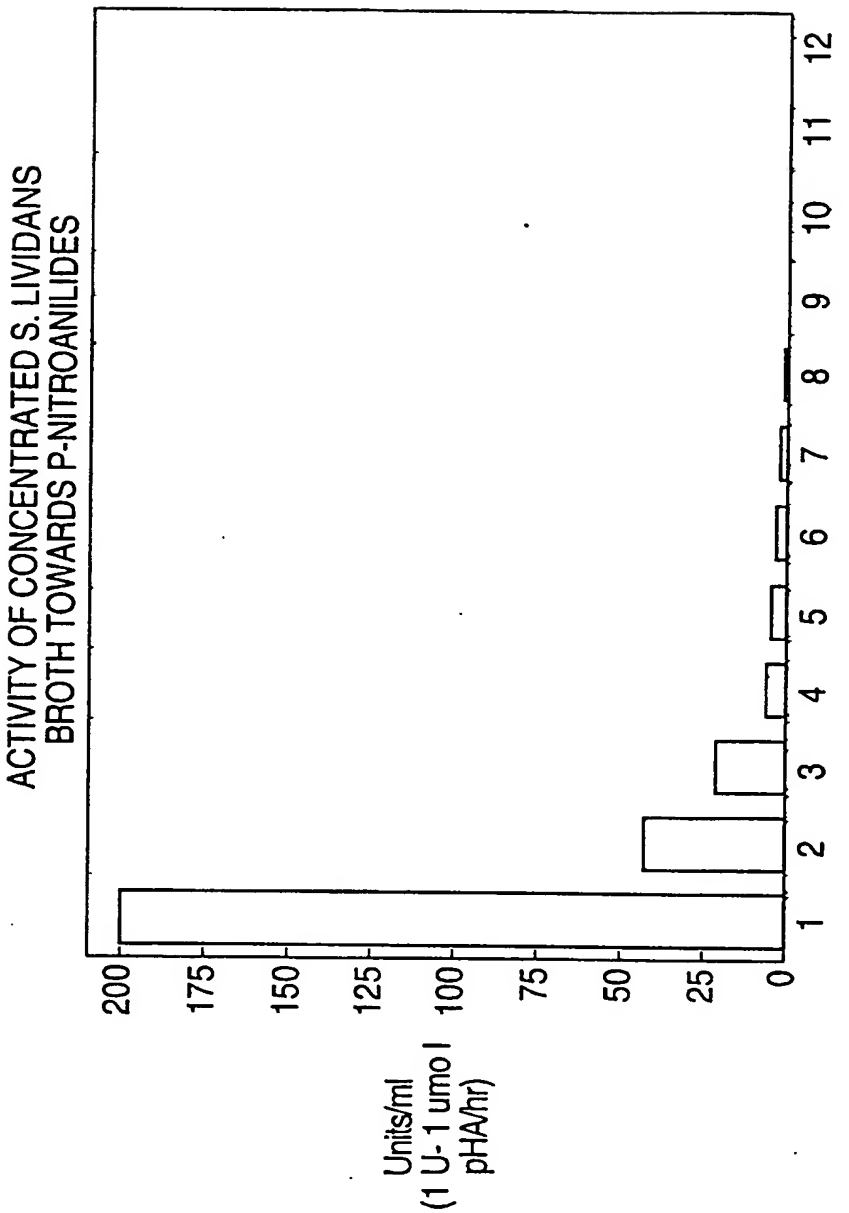

ACTIVITY OF CONCENTRATED S. LIVIDANS
BROTH TOWARDS P-NITROANILIDES
Units/ml
(1 U- 1 umol
pHA/hr)
200
175
150
125
100
75
50
25
0
1
2
3
4
5
6
7
8
9
10
11
12

FIG. 5(A)

```
GGCGGGGACC GGCCGACGGC CCCGCCGAAC GAACGCCCTT CTCCGTTTAT CGGATTGGCA   60
AAGAAGTAGC ACTGGCCCTG TTCTCAGGAA ACCCACAGCG GCGAGGATCC CCGTACTTGT  120
CGCGAACACG TACGGGGAGG GCCAC TTG AGG AAG AGC AGC ATA CGG CGG AGG     172
                           fMet Arg Lys Ser Ser Ile Arg Arg Arg
                                              -35

GCG ACC GCC TTC GGC ACG GCC GGA GCA CTG GTC ACC GCC ACG CTG ATC     220
Ala Thr Ala Phe Gly Thr Ala Gly Ala Leu Val Thr Ala Thr Leu Ile
-30                 -25                 -20                 -15

GCC GGC GCC GTC TCG GCA CCC GCC GCG AGC GCC GCC CCG GCC GAC GGC     268
Ala Gly Ala Val Ser Ala Pro Ala Ala Ser Ala Ala Pro Ala Asp Gly
                -10                 -5                  1

CAC GGG CAC GGG CGG AGC TGG GAC CGG GAG GCG CGC GGT GCC GCC ATC     316
His Gly His Gly Arg Ser Trp Asp Arg Glu Ala Arg Gly Ala Ala Ile
        5                   10                  15

GCC GCC GCC CGC GCC GCC CGG GCG GGC ATC GAC TGG GAG GAC TGC GCA     364
Ala Ala Ala Arg Ala Ala Arg Ala Gly Ile Asp Trp Glu Asp Cys Ala
    20                  25                  30

GCC GAC TGG AAC CTG CCC AAG CCC ATC CAG TGC GGC TAC GTC ACG GTG     412
Ala Asp Trp Asn Leu Pro Lys Pro Ile Gln Cys Gly Tyr Val Thr Val
35                  40                  45                  50

CCG ATG GAC TAC GCC AAG CCG TAC GGC AAG CAG ATC AGG CTC GCC GTC     460
Pro Met Asp Tyr Ala Lys Pro Tyr Gly Lys Gln Ile Arg Leu Ala Val
                55                  60                  65
```

FIG. 5(B)

```
GAC CGC ATC GGC AAC ACC GGA ACC AGG AGC GAG CGC CAG GGC GCC CTG      508
Asp Arg Ile Gly Asn Thr Gly Thr Arg Ser Glu Arg Gln Gly Ala Leu
            70                  75                  80

ATC TAC AAC CCC GGC GGT CCC GGC GGC TCC GGC CTG CGT TTC CCG GCC      556
Ile Tyr Asn Pro Gly Gly Pro Gly Gly Ser Gly Leu Arg Phe Pro Ala
        85                  90                  95

CGC GTC ACG AAC AAG AGC GCG GTC TGG GCC AAC ACG GCC AAG GCC TAC      604
Arg Val Thr Asn Lys Ser Ala Val Trp Ala Asn Thr Ala Lys Ala Tyr
    100                 105                 110

GAC TTC GTC GGC TTC GAC CCG CGC GGC GTC GGC CAC TCC GCG CCC ATC      652
Asp Phe Val Gly Phe Asp Pro Arg Gly Val Gly His Ser Ala Pro Ile
115                 120                 125                 130

TCC TGC GTC GAC CCG CAG GAG TTC GTC AAG GCA CCC AAG GCC GAC CCC      700
Ser Cys Val Asp Pro Gln Glu Phe Val Lys Ala Pro Lys Ala Asp Pro
                135                 140                 145

GTG CCC GGC TCC GAG GCC GAC AAG CGC GCC CAG CGC AAG CTC GCC CGC      748
Val Pro Gly Ser Glu Ala Asp Lys Arg Ala Gln Arg Lys Leu Ala Arg
            150                 155                 160

GAG TAC GCC GAG GGC TGC TTC GAG CGC AGC GGC GAG ATG CTC CCG CAC      796
Glu Tyr Ala Glu Gly Cys Phe Glu Arg Ser Gly Glu Met Leu Pro His
        165                 170                 175
```

FIG. 5(C)

```
ATG ACC ACG CCG AAC ACC GCG CGC GAC CTC GAC GTC ATC CGC GCC GCC     844
Met Thr Thr Pro Asn Thr Ala Arg Asp Leu Asp Val Ile Arg Ala Ala
    180                 185                 190

CTC GGC GAG AAG AAG CTC AAC TAC CTC GGC GTC TCC TAC GGC ACC TAC     892
Leu Gly Glu Lys Lys Leu Asn Tyr Leu Gly Val Ser Tyr Gly Thr Tyr
195                 200                 205                 210

CTC GGC GCC GTC TAC GGC ACC CTC TTC CCG GAC CAC GTC CGC CGC ATG     940
Leu Gly Ala Val Tyr Gly Thr Leu Phe Pro Asp His Val Arg Arg Met
                215                 220                 225

GTC GTC GAC AGC GTC GTC AAC CCG TCC CGC GAC AAG ATC TGG TAC CAG     988
Val Val Asp Ser Val Val Asn Pro Ser Arg Asp Lys Ile Trp Tyr Gln
            230                 235                 240

GCC AAC CTG GAC CAG GAC GTC GCC TTC GAG GGC CGC TGG AAG GAC TGG    1036
Ala Asn Leu Asp Gln Asp Val Ala Phe Glu Gly Arg Trp Lys Asp Trp
        245                 250                 255

CAG GAC TGG GTC GCC GCG AAC GAC GCC GCC TAC CAC CTC GGC GAC ACC    1084
Gln Asp Trp Val Ala Ala Asn Asp Ala Ala Tyr His Leu Gly Asp Thr
    260                 265                 270

CGC GCC GAG GTC CAG GAC CAG TGG CTG AAG CTG CGC GCC GCC GCC GCG    1132
Arg Ala Glu Val Gln Asp Gln Trp Leu Lys Leu Arg Ala Ala Ala Ala
275                 280                 285                 290

AAG AAG CCG CTG GGC GGC GTC GTC GGA CCG GCC GAG CTG ATC TCC TTC    1180
Lys Lys Pro Leu Gly Gly Val Val Gly Pro Ala Glu Leu Ile Ser Phe
                295                 300                 305
```

FIG. 5(D)

```
TTC CAG AGC GCC CCG TAC TAC GAC TCC GCC TGG GCG CCG ACC GCG GAG   1228
Phe Gln Ser Ala Pro Tyr Tyr Asp Ser Ala Trp Ala Pro Thr Ala Glu
            310                 315                 320

ATC TTC AGC AAG TAC GTC GCC GGC GAC ACC CAG GCG CTC GTC GAC GCC   1276
Ile Phe Ser Lys Tyr Val Ala Gly Asp Thr Gln Ala Leu Val Asp Ala
        325                 330                 335

GCC GCA CCC GAC CTG TCC GAC ACC GCG GGC AAC GCC TCC GCG GAG AAC   1324
Ala Ala Pro Asp Leu Ser Asp Thr Ala Gly Asn Ala Ser Ala Glu Asn
    340                 345                 350

GGC AAC GCC GTC TAC ACG GCC GTC GAG TGC ACC GAC GCC AAG TGG CCC   1372
Gly Asn Ala Val Tyr Thr Ala Val Glu Cys Thr Asp Ala Lys Trp Pro
355                 360                 365                 370

GCC AAC TGG CGC ACC TGG GAC CGG GAC AAC ACC CGG CTC CAC CGC GAC   1420
Ala Asn Trp Arg Thr Trp Asp Arg Asp Asn Thr Arg Leu His Arg Asp
                375                 380                 385

CAC CCG TTC ATG ACC TGG GCC AAC GCC TGG ATG AAC CTG CCC TGT GCC   1468
His Pro Phe Met Thr Trp Ala Asn Ala Trp Met Asn Leu Pro Cys Ala
            390                 395                 400

ACC TGG CCG GTC AAG CAG CAG ACC CCG CTG AAC GTG AAG ACC GGC AAG   1516
Thr Trp Pro Val Lys Gln Gln Thr Pro Leu Asn Val Lys Thr Gly Lys
        405                 410                 415
```

FIG. 5(E)

```
GGA CTT CCG CCG GTG CTG ATC GTC CAG TCC GAG CGT GAC GCC GCC ACC   1564
Gly Leu Pro Pro Val Leu Ile Val Gln Ser Glu Arg Asp Ala Ala Thr
    420                 425                 430

CCG TAC GAG GGC GCC GTC GAA CTG CAC CAG CGG TTC CGG GGA TCC CGC   1612
Pro Tyr Glu Gly Ala Val Glu Leu His Gln Arg Phe Arg Gly Ser Arg
435                 440                 445                 450

CTG ATC ACC GAG CGG GAC GCC GGC TCC CAC GGC GTC ACC GGC CTG GTC   1660
Leu Ile Thr Glu Arg Asp Ala Gly Ser His Gly Val Thr Gly Leu Val
                455                 460                 465

AAC CCG TGC ATC AAC GAC CGG GTC GAC ACC TAC CTG CTC ACC GGC AGG   1708
Asn Pro Cys Ile Asn Asp Arg Val Asp Thr Tyr Leu Leu Thr Gly Arg
            470                 475                 480

ACG GAC GCC CGC GAC GTG ACC TGC GCG CCG CAC GCC ACG CCC AGG CCG   1756
Thr Asp Ala Arg Asp Val Thr Cys Ala Pro His Ala Thr Pro Arg Pro
        485                 490                 500

TAA CCCGGGCTCA GGCCAAGCGG GGGGAGGGGG CGACCGGTCC GACCGGCCGC       1809
End
CCCCTCCCCC CACCTGTCGC TACCGTCCCT CGGCCCAGGC GTCCTCCGCC GCGTAGTCGA 1869
AGAGGTCGCC GTACGCCTTG AACATCTTCG GGTAGGCCT                       1908
```

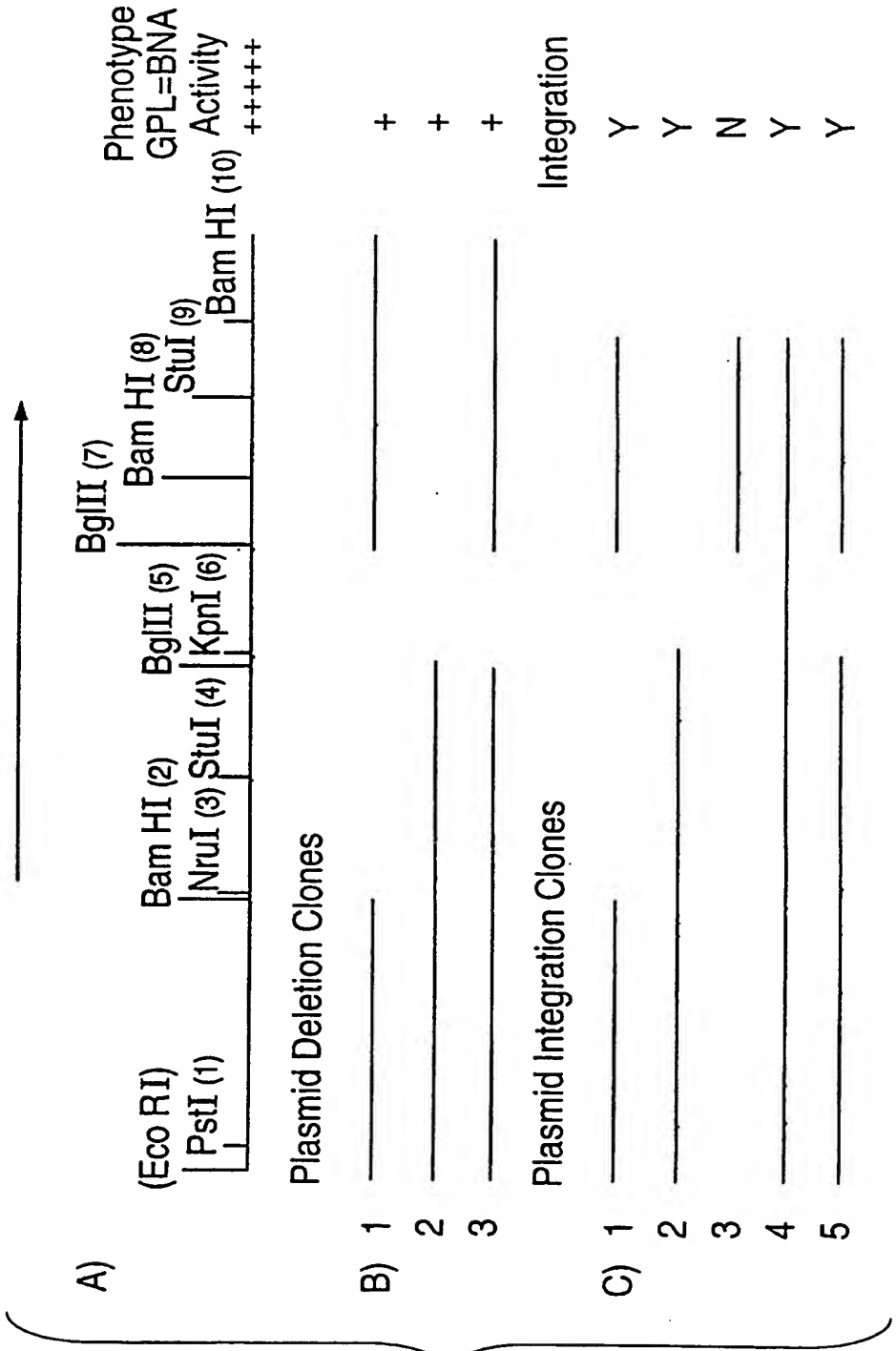

A)
(Eco RI)
PstI (1)
Bam HI (2)
NruI (3)
StuI (4)
BglII (5)
KpnI (6)
BglII (7)
Bam HI (8)
StuI (9)
Bam HI (10)
Plasmid Deletion Clones
Phenotype
GPL=BNA
Activity
B) 1 +
2 +
3 +
Plasmid Integration Clones
Integration
C) 1 Y
2 Y
3 N
4 Y
5 Y

FIG. 6

FIG. 7

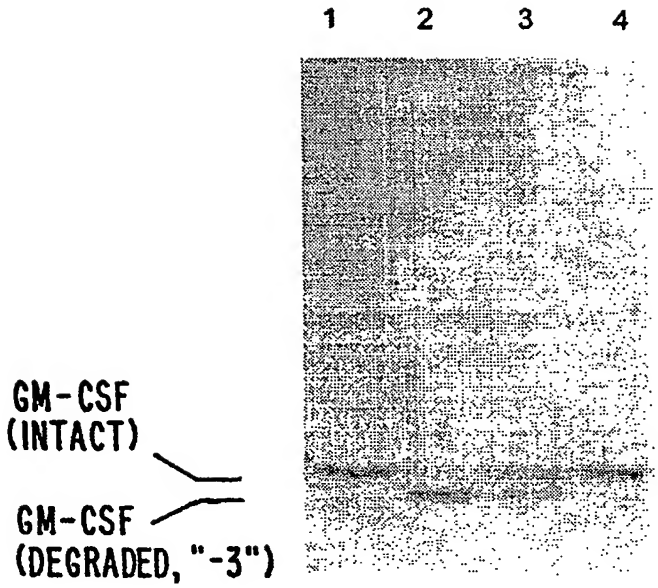


FIG. 8

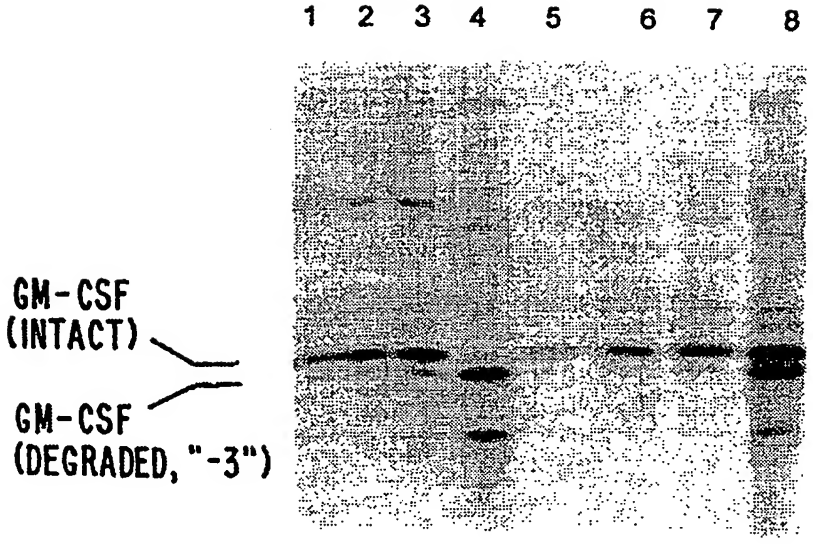
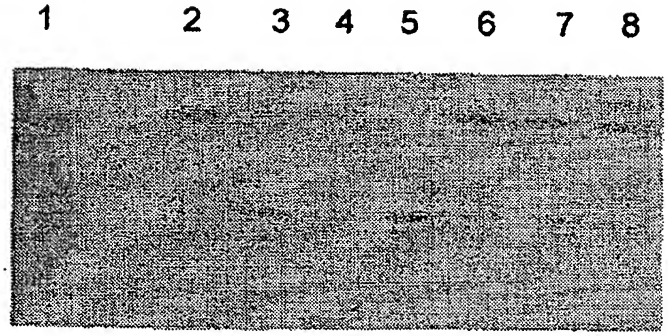


FIG. 9



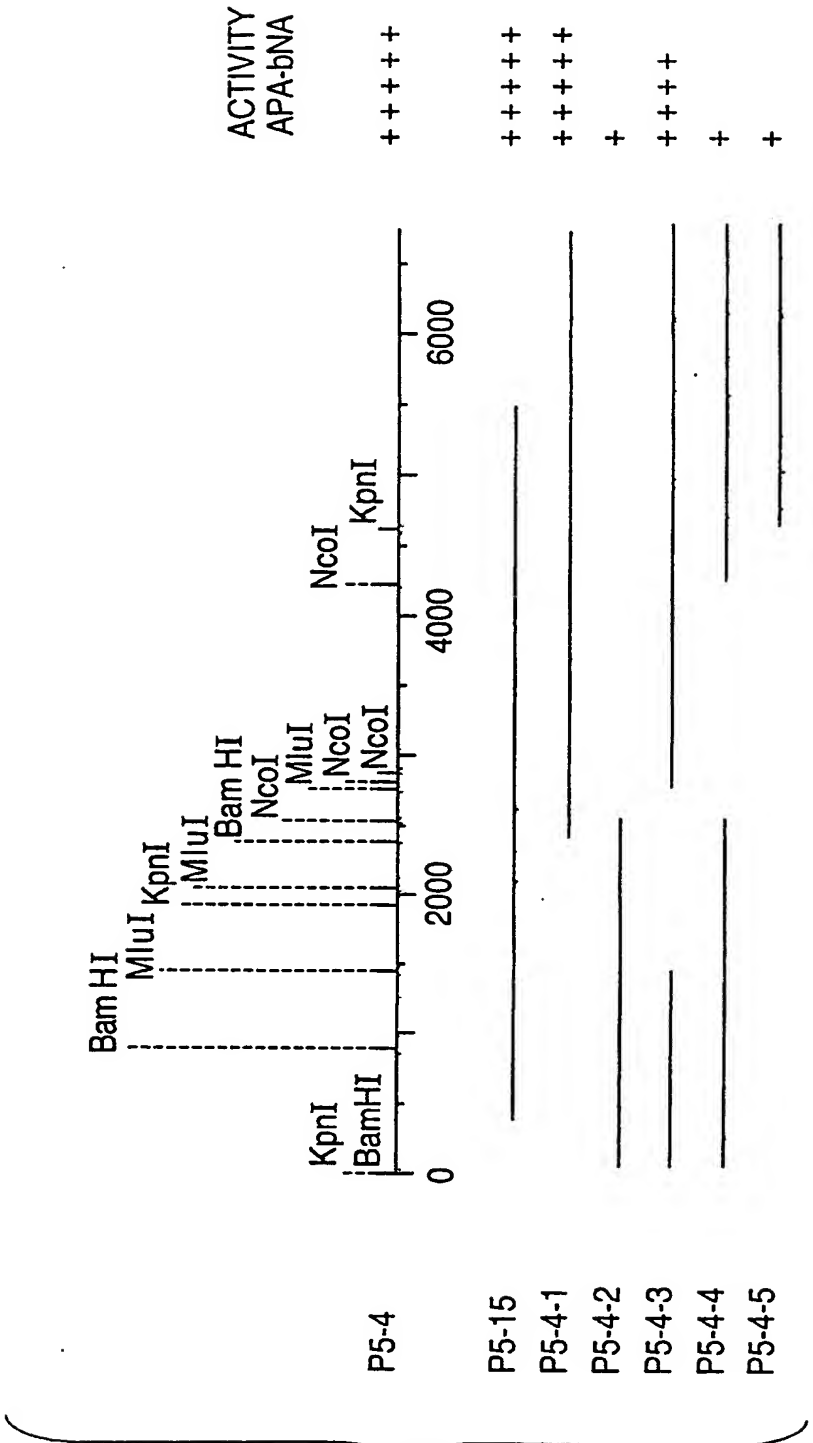

ACTIVITY
APA-bNA
+++++
+++++
+++++
+
++++
+
+
KpnI
BamHI
Bam HI
MluI
KpnI
MluI
Bam HI
NcoI
MluI
NcoI
NcoI
NcoI
KpnI
0
2000
4000
6000
P5-4
P5-15
P5-4-1
P5-4-2
P5-4-3
P5-4-4
P5-4-5

FIG. 10

FIG. 11(A)

```
GGTACCAGGC GACGAAGGCG ACGGTCAGCG GGAACGCGAA GGAACGGAAG GAGCGGCGCA    60
GTTCGGCGAA CTCGGCGCTC TGCTGCACTT CGGAGAACTC CTCGGCGGAG GGGAGGCGGT   120
GCTCCTCTTG CGAGGGGGGC TCCTCTTTGG AGGGGGGCGG TGCGTCGGGT GGCCACGGAG   180
TCTCCTCGTA CGACGGACAT GACGGCTTGG ACCTCGGTGT TCTCGCAGGG GGCTGATCGT   240
GCTCGGGCTC CCTGTCCAAC GACACGGCGC CCCGCGGGGC CCGGTTCAAC ACCCGTGGCA   300
CTTTCCGAAG TCGTCCTCGG CGGGTCATTG CTGGCCAGGG ACTTCGGGGG ATAGCTTCAC   360
CCTGCACCAC TACGTCATGT ACCTGCCCGG CCCGTTTCAC CCGTGCCCGG GCAGGTGCTG   420
TTTGCCGGAT GATGTGGAGA CCCCATGGAT CATCTGCGCT TCCCGCGCGA CCCGCGCTCC   480
AGACGCGGGC TCGTTTCCCG AGCTTTCCCG ACGGACTGGA GACATCACGC  ATG ACC    536
                                                        fMet Thr

GCT CCC CTC TCG CGT CAC CGC CGT GCC CTC GCG ATT CCG GCG GGC CTG     584
Ala Pro Leu Ser Arg His Arg Arg Ala Leu Ala Ile Pro Ala Gly Leu
       -120                 -115                 -110

GCC GTG GCC GCG TCG CTC GCG TTC CTG CCG GGC ACC CCG GCC GCC GCG     632
Ala Val Ala Ala Ser Leu Ala Phe Leu Pro Gly Thr Pro Ala Ala Ala
   -105                 -100                 -95

ACC CCC GCG GCC GAG GCC GCG CCC TCG ACG GCG GCG GAC GCG ACC TCG     680
Thr Pro Ala Ala Glu Ala Ala Pro Ser Thr Ala Ala Asp Ala Thr Ser
-90                  -85                  -80                  -75

CTC AGC TAC GTC GTC AAC GTC GCC TCC GGG CAC CGT CCT TCG GCC ACC     728
Leu Ser Tyr Val Val Asn Val Ala Ser Gly His Arg Pro Ser Ala Thr
                -70                  -65                  -60
```

FIG. 11(B)

```
GTG CGG CGG GCG ATA GCC AAG GCG GGC GGC ACG ATC GTC ACG TCG TAC     776
Val Arg Arg Ala Ile Ala Lys Ala Gly Gly Thr Ile Val Thr Ser Tyr
            -55                 -50                 -45

GAC CGG ATC GGC GTG ATC GTC GTC CAC TCC GCC AAC CCC GAC TTC GCC     824
Asp Arg Ile Gly Val Ile Val Val His Ser Ala Asn Pro Asp Phe Ala
        -40                 -35                 -30

AAG ACC GTG CGC AAG GTG CGC GGC GTG CAG TCG GCC GGT GCC ACC CGC     872
Lys Thr Val Arg Lys Val Arg Gly Val Gln Ser Ala Gly Ala Thr Arg
    -25                 -20                 -15

ACC GCG CCA CTG CCC TCG GCC GCC ACC ACC GAC ACG GGC GCG CCG CAG     920
Thr Ala Pro Leu Pro Ser Ala Ala Thr Thr Asp Thr Gly Ala Pro Gln
-10                 -5                  1                   5

GTG CTC GGC GGC GAG GAC CTG GCC GCC GCC AAG GCC GCC TCC GCG AAG     968
Val Leu Gly Gly Glu Asp Leu Ala Ala Ala Lys Ala Ala Ser Ala Lys
            10                  15                  20

GCC GAG GGC CAG GAC CCG CTG GAG TCG CTC CAG TGG GAC CTG CCC GCC    1016
Ala Glu Gly Gln Asp Pro Leu Glu Ser Leu Gln Trp Asp Leu Pro Ala
        25                  30                  35
```

FIG. 11(C)

```
ATC AAG GCG GAC AAG GCG CAC GAG AAG TCG CTG GGC AGC AGG AAG GTG   1064
Ile Lys Ala Asp Lys Ala His Glu Lys Ser Leu Gly Ser Arg Lys Val
     40                  45                  50

ACC GTC GCC GTC ATC GAC ACC GGC GTC GAC GAC ACC CAC CCG GAC ATC   1112
Thr Val Ala Val Ile Asp Thr Gly Val Asp Asp Thr His Pro Asp Ile
 55                  60                  65                  70

GCC CCG AAC TTC GAC CGG CAG GCG TCC GTC AAC TGT GTG GCG GGC AAG   1160
Ala Pro Asn Phe Asp Arg Gln Ala Ser Val Asn Cys Val Ala Gly Lys
                 75                  80                  85

CCG GAC ACC GCC GAC GGG GCC TGG CGG CCG AGC GCG GCG GAG AGC CCG   1208
Pro Asp Thr Ala Asp Gly Ala Trp Arg Pro Ser Ala Ala Glu Ser Pro
             90                  95                 100

CAC GGC ACC CAC GTG GCC GGG GAG ATA GCC GCC GCC AAG AAC GGC GTC   1256
His Gly Thr His Val Ala Gly Glu Ile Ala Ala Ala Lys Asn Gly Val
        105                 110                 115

GGC ATG ACC GGC GTG GCA CCC GGG GTG AAG GTG GCC GGC ATC AAG GTC   1304
Gly Met Thr Gly Val Ala Pro Gly Val Lys Val Ala Gly Ile Lys Val
    120                 125                 130
```

FIG. 11(D)

```
TCC AAC CCC GAC GGC TTC TTC TAC ACC GAG GCC GTG GTC TGC GGC TTC   1352
Ser Asn Pro Asp Gly Phe Phe Tyr Thr Glu Ala Val Val Cys Gly Phe
135                 140                 145                 150

ATG TGG GCG GCC GAG CAC GGC GTC GAC GTG ACC AAC AAC AGC TAT TAC   1400
Met Trp Ala Ala Glu His Gly Val Asp Val Thr Asn Asn Ser Tyr Tyr
                155                 160                 165

ACC GAC CCG TGG TAC TTC AAC TGC AAG GAC GAG CCC GAC CAG AAG GCG   1448
Thr Asp Pro Trp Tyr Phe Asn Cys Lys Asp Asp Pro Asp Gln Lys Ala
            170                 175                 180

CTC GTC GAG GCC GTC TCG CGG GCC TCC CGG TAC GCG GAG AAG AAG GGC   1496
Leu Val Glu Ala Val Ser Arg Ala Ser Arg Tyr Ala Glu Lys Lys Gly
        185                 190                 195

GCG GTC AAC GTC GCC GCG GCC GGC AAC GAG AAC TAC GAC CTC ACC TCC   1544
Ala Val Asn Val Ala Ala Ala Gly Asn Glu Asn Tyr Asp Leu Thr Ser
    200                 205                 210

GAC GAG ATC ACC GAC CCG TCC TCG CCC AAC GAC ACC ACG CCC GGC GAC   1592
Asp Glu Ile Thr Asp Pro Ser Ser Pro Asn Asp Thr Thr Pro Gly Asp
215                 220                 225                 230
```

FIG. 11(E)

```
CGG ACC GTC GAC CCG TCG AAG TGC CTG GAC ATC CCG ACC CAG CTG CCG   1640
Arg Thr Val Asp Pro Ser Lys Cys Leu Asp Ile Pro Thr Gln Leu Pro
                235                 240                 245

GGT GTC GTG ACG GTC GCG GCG ACC GGT GCG AAG GGC CTC AAG TCG TCC   1688
Gly Val Val Thr Val Ala Ala Thr Gly Ala Lys Gly Leu Lys Ser Ser
            250                 255                 260

TTC TCC AAC CAC GGG CTG GGC GTC ATC GAC ATC GCC GCG CCC GGC GGC   1736
Phe Ser Asn His Gly Leu Gly Val Ile Asp Ile Ala Ala Pro Gly Gly
        265                 270                 275

GAC TCG ACG GCC TAC CAG ACC CCG GAG CCG CCC GCC ACG AGC GGC CTG   1784
Asp Ser Thr Ala Tyr Gln Thr Pro Glu Pro Pro Ala Thr Ser Gly Leu
    280                 285                 290

ATC CTG GGC ACG CTG CCC GGC GGC AAG TGG GGC TAC ATG GCC GGT ACG   1832
Ile Leu Gly Thr Leu Pro Gly Gly Lys Trp Gly Tyr Met Ala Gly Thr
295                 300                 305                 310

TCC ATG GCC TCC CCG CAC GTC GCG GGC GTC GCC GCC CTC ATC AAG TCG   1880
Ser Met Ala Ser Pro His Val Ala Gly Val Ala Ala Leu Ile Lys Ser
                315                 320                 325
```

FIG. 11(F)

```
ACG CAC CCG CAC GCC TCC CCC GCC ATG GTG AAG GCG CTG CTG TAC GCC   1928
Thr His Pro His Ala Ser Pro Ala Met Val Lys Ala Leu Leu Tyr Ala
            330                 335                 340

GAG GCC GAC GCC ACG GCG TGC ACC AAG CCG TAC GAC ATC GAC GGC GAC   1976
Glu Ala Asp Ala Thr Ala Cys Thr Lys Pro Tyr Asp Ile Asp Gly Asp
        345                 350                 355

GGC AAG GTC GAC GCG GTG TGC GAG GGC CCG AAG AAC CGC AAC GGC TTC   2024
Gly Lys Val Asp Ala Val Cys Glu Gly Pro Lys Asn Arg Asn Gly Phe
    360                 365                 370

TAC GGC TGG GGC ATG GCC GAC GCG CTG GAC GCG GTG ACC TGG TAG CCGGT
Tyr Gly Trp Gly Met Ala Asp Ala Leu Asp Ala Val Thr Trp ter
375                 380                 385

ACGCGTACCC GGTGCGTGAG GCGGGGGCGG CGGTCCGGTT CCCGTCCGGT CCGCCGCCCC 2074
CGTCGTCGTC GTCGTACGAC AGTATCTTCG CCATGGACAC TTACGAGGAT CC          2185
```

FIG. 12(A)

```
          10v        20v        30v        40v        50v        60v
1   MTAPLSRHRRALAIPAGLAVAASLAFLPGTPAAATPAAEAAPSTAADATSLSYVVNVASGH
2                                                    :..:   :  A :
                                                    MRGKKVWISLLFALAL
                                                              10^
          70v        80v        90v       100v       110v       120v
1   RPSATVRRAIAKAGGTIVTSYDRIGVIVVHSANPDFAKTVRKVRGVQSAGATRTAPLPSAA
      : :  .: : .::.  .: .:  V   :: :.  A . :.V : :::. :     .AA
2   IFTMAFGSTSSAQAAGKSNGEKKYIVGFKQTMSTMSAAKKKDVISEKGGKVQKQFKYVDAA
     20^        30^        40^        50^        60^        70^
        130v       140v       150v       160v       170v       180v
1   TTDTGAPQVLGGEDLAAAKAASAKAEGQDPLESLQWDLPAIKADKAHEKSLGSRKVTVAVI
    ::. .. .V . .. ::.  .... ::.  :S:.::::.IKA   H.:: .:.:V.VAVI
2   SATLNEKAVKELKKDPSVAYVEEDHVAHAYAQSVPYGVSQIKAPALHSQGYTGSNVKVAVI
    80^        90^       100^       110^       120^       130^
       190v       200v       210v       220v       230v       240v
1   DTGVDDTHPDIAPNFDRQASVNCVAGKPDTADGAWRPSAAESPHGTHVAGEIAAAKNGVGM
    D:G:D.:HPD:            VAG :. . :...P ..::::HGTHVAG.:AA :N::G:
2   DSGIDSSHPDL----------KVAGGASMVPSETNPFQDNNSHGTHVAGTVAALNNSIGV
  140^                   150^       160^       170^       180^
      250v       260v       270v       280v       290v       300v
1   TGRWHPGVKVAGIKVSNPDGFFYTEAVVCGFMWAAEHGVDVTNNSYYTDPWYFNCKDDPDQ
```

FIG. 12(B)

```
     G   P:..: ::KV .:DG    . :: G: WA .:.:DV.N S  . :
2    LGV-APSASLYAVKVLGADGSGQYSWIINGIEWAIANNMDVINMSLGGPSGS---------
    190^       200^      210^      220^      230^
     310v      320v      330v      340v      350v            360v
1    KALVEAVSRASRYAEKKGAVNVAAAGNENYDLTSDEITDPSSPNDTTPGDRTVD-----PS
     AL  AV  A       G V VAAAGNE          T  SS     PG
2    AALKAAVDKA---V-ASGVVVVAAAGNEG---------TSGSSSTVGYPG-KYPSVIAVGAV
   240^          250^      260^              270^       280^
        370v      380v      390v      400v      410v      420v
1    KCLDIPTQLPGVVTVAATGAKGLKSSFSNHGLGVIDIAAPGGDSTAYQTPEPPATSGL-IL
                               F   G    D                    G: I
2    DSSN-------------------RASFSSVG-PELD-----------------VMAPGVSIQ
     290^                          300^                      310^
        430v      440v      450v      460v      470v      480v
1    GTLPGGKWGYMAGTSMASPHVAGVAALIKSTHPHASPAMVKALLYAEADATACTKPYDIDG
     :TLPG.K.G  .GTSMASPHVAG.AALI S.HP: : : V:: L ..:.  :    Y: :
2    STLPGNKYGAYNGTSMASPHVAGAAALILSKHPNWTNTQVRSSLENTTTKLGDSFYYGKGL
        320^      330^      340^      350^      360^      370^
       490v      500v      510v
1    DGKVDAVCEGPKNRNGFYGWGMADALDAVTW
     .  .A.
2    INVQAAAQ
       380^
```

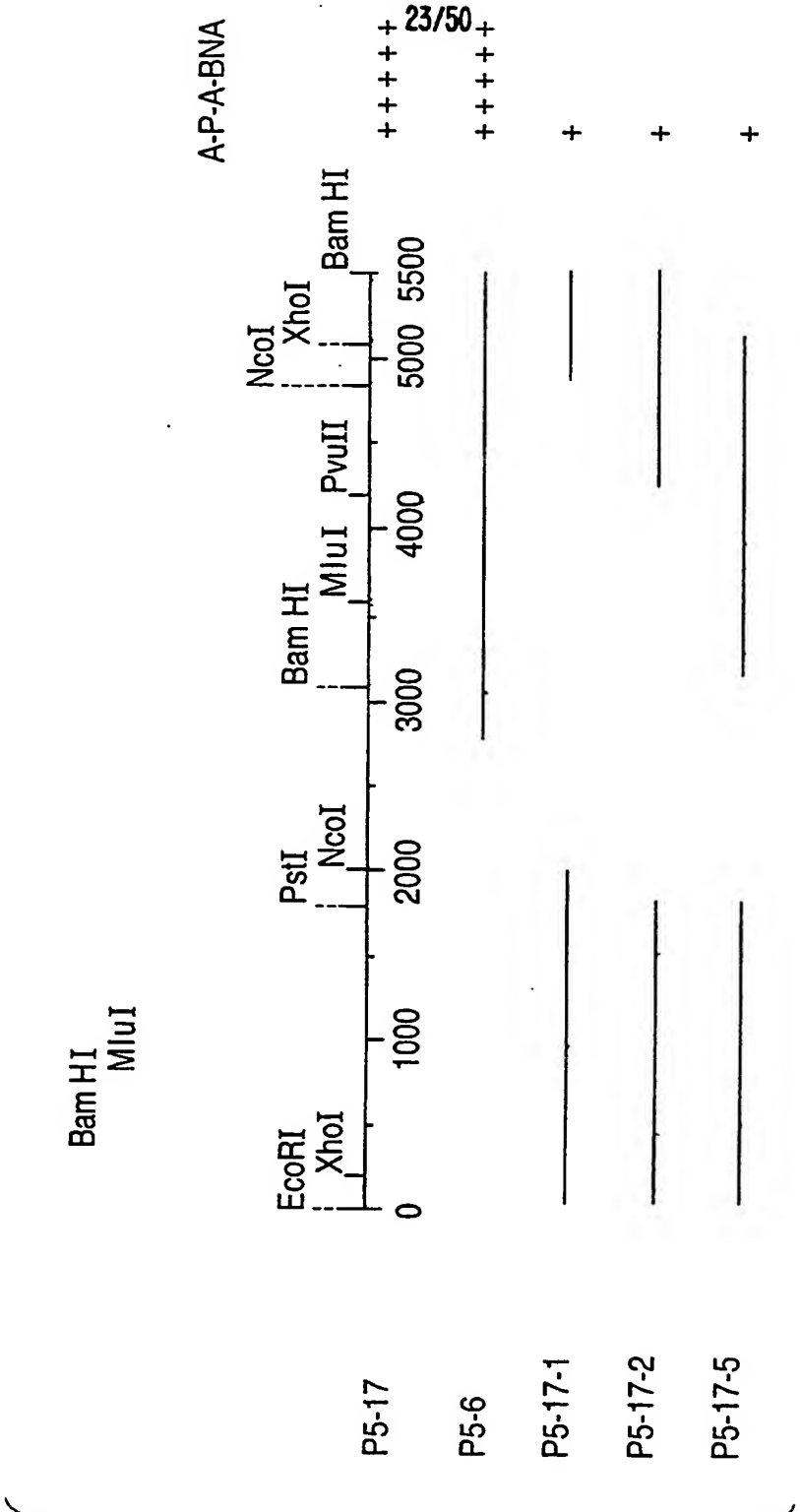

A-P-A-BNA
23/50
+ + + + +
+ + + + +
+
+
+
Bam HI
MluI
EcoRI
XhoI
PstI
NcoI
Bam HI
MluI
PvuII
NcoI
XhoI
Bam HI
0
1000
2000
3000
4000
5000
5500
P5-17
P5-6
P5-17-1
P5-17-2
P5-17-5

FIG. 13

FIG. 14(A)

```
CCCGGGCCCG CGTCGGAGTC ATGACCGGTT GACGCCGTAA CACGTACGGG GCACGCGCAC      60
CACGCACCGC AACTGCTTCG TCGCGGAGAG TTACGCTCGC TGA ATG GAC ACA AGG      115
                                                Met Asp Thr Arg
                                                    -45

CGC ACT CAC CGC AGG ACC CGC ACC GGC GGC ACC CGT TTC CGG GCC ACG      163
Arg Thr His Arg Arg Thr Arg Thr Gly Gly Thr Arg Phe Arg Ala Thr
        -40                 -35                 -30

CTG CTC ACC GCC GCG CTG CTC GCC ACC GCC TGC TCG GCC GGG GGC GCG      211
Leu Leu Thr Ala Ala Leu Leu Ala Thr Ala Cys Ser Ala Gly Gly Ala
    -25                 -20                 -15

TCG ACG TCC GCC GGA TCC CCC GCG GCC AAG GCG GCC GGC GCG ACG GAG      259
Ser Thr Ser Ala Gly Ser Pro Ala Ala Lys Ala Ala Gly Ala Thr Glu
-10                 -5                  1                   5

GCG GCC ACG GCG ACC CTG ACC CCC CTG CCG AAG GCC ACG CCC GCC GAG      307
Ala Ala Thr Ala Thr Leu Thr Pro Leu Pro Lys Ala Thr Pro Ala Glu
            10                  15                  20

CTG TCC CCG TAC TAC GAG CAG AAG CTC GGC TGG CGC GAC TGC GGC GTC      355
Leu Ser Pro Tyr Tyr Glu Gln Lys Leu Gly Trp Arg Asp Cys Gly Val
        25                  30                  35
```

FIG. 14(B)

```
CCG GGC TTC CAG TGC GCC ACC ATG AAG GCC CCG CTC GAC TAC GCC AAG          403
Pro Gly Phe Gln Cys Ala Thr Met Lys Ala Pro Leu Asp Tyr Ala Lys
          40                  45                  50

CCC GCC GAC GGC GAC GTC CGG CTC GCG GTG GCC CGC AAG AAG GCC ACG          451
Pro Ala Asp Gly Asp Val Arg Leu Ala Val Ala Arg Lys Lys Ala Thr
     55                  60                  65

GGG CCG GGC AAG CGC CTC GGC TCG CTG CTG GTC AAC CCG GGC GGA CCG          499
Gly Pro Gly Lys Arg Leu Gly Ser Leu Leu Val Asn Pro Gly Gly Pro
 70                  75                  80                  85

GGC GGC TCG GCG ATC GGC TAC CTC CAG CAG TAC GCG GGC ATC GGC TAC          547
Gly Gly Ser Ala Ile Gly Tyr Leu Gln Gln Tyr Ala Gly Ile Gly Tyr
                 90                  95                 100

CCG GCG AAG GTC CGC GCC CAG TAC GAC ATG GTG GCG GTC GAC CCC CGG          595
Pro Ala Lys Val Arg Ala Gln Tyr Asp Met Val Ala Val Asp Pro Arg
            105                 110                 115

GGC GTG GCC CGC AGT GAA CCC GTC GAG TGC CTG GAC GGG CGC GAG ATG          643
Gly Val Ala Arg Ser Glu Pro Val Glu Cys Leu Asp Gly Arg Glu Met
        120                 125                 130

GAC GCG TAC ACG CGC ACC GAC GTC ACC CCG GAC GAC GCG GGC GAG ACG          691
Asp Ala Tyr Thr Arg Thr Asp Val Thr Pro Asp Asp Ala Gly Glu Thr
    135                 140                 145
```

FIG. 14(C)

```
GAC GAG CTG GTC GAC GCC TAC AAG GAG TTC GCC GAG GGC TGC GGG GCG          739
Asp Glu Leu Val Asp Ala Tyr Lys Glu Phe Ala Glu Gly Cys Gly Ala
150                 155                 160                 165

GAC GCG CCG AAG CTG CTG CGC CAC GTC TCC ACG GTC GAG GCG GCA CGC          789
Asp Ala Pro Lys Leu Leu Arg His Val Ser Thr Val Glu Ala Ala Arg
                170                 175                 180

GAC ATG GAC GTC CTG CGC GCG GTG CTG GGC GAC GAG AAG CTG ACC TAC          835
Asp Met Asp Val Leu Arg Ala Val Leu Gly Asp Glu Lys Leu Thr Tyr
            185                 190                 195

GTG GGA GCG TCG TAC GGC ACC TTC CTG GGC GCG ACC TAC GCC GGT CTG          883
Val Gly Ala Ser Tyr Gly Thr Phe Leu Gly Ala Thr Tyr Ala Gly Leu
        200                 205                 210

TTC CCC GAC CGG ACG GGC CGC CTG GTC CTG GAC GGC GCG ATG GAC CCC          931
Phe Pro Asp Arg Thr Gly Arg Leu Val Leu Asp Gly Ala Met Asp Pro
    215                 220                 225

TCG CTG CCC GCC CGC CGC CTG AAC CTG GAG CAG ACG GAG GGC TTC GAG          979
Ser Leu Pro Ala Arg Arg Leu Asn Leu Glu Gln Thr Glu Gly Phe Glu
230                 235                 240                 245

ACG GCG TTC CAG TCC TTC GCG AAG GAC TGC GTG AAG CAG CCG GAC TGC         1027
Thr Ala Phe Gln Ser Phe Ala Lys Asp Cys Val Lys Gln Pro Asp Cys
                250                 255                 260
```

FIG. 14(D)

```
CCC CTC GGC GAC AAG GAC ACC ACC CCC GAC CAG GTC GGC AAG AAC CTC          1075
Pro Leu Gly Asp Lys Asp Thr Thr Pro Asp Gln Val Gly Lys Asn Leu
            265                 270                 275

AAG TCC TTC TTC GAC GAC CTG GAC GCG AAG CCC CTG CCC GCC GGC GAC          1123
Lys Ser Phe Phe Asp Asp Leu Asp Ala Lys Pro Leu Pro Ala Gly Asp
        280                 285                 290

GCC GAC GGC CGC AAG CTC ACC GAA TCC CTC GCC ACC ACC GGC GTG ATC          1171
Ala Asp Gly Arg Lys Leu Thr Glu Ser Leu Ala Thr Thr Gly Val Ile
    295                 300                 305

GCC GCG ATG TAC GAC GAG GGC GCC TGG CAG CAG CTG CGC GAG TCC CTC          1219
Ala Ala Met Tyr Asp Glu Gly Ala Trp Gln Gln Leu Arg Glu Ser Leu
310                 315                 320                 325

ACC TCG GCG ATC AAG GAG AAG GAC GGT GCG GGC CTG CTG ATC CTC TCC          1267
Thr Ser Ala Ile Lys Glu Lys Asp Gly Ala Gly Leu Leu Ile Leu Ser
                330                 335                 340

GAC AGC TAC TAC GAG CGC GAG GCC GAC GGC GGC TAC AGC AAC CTG ATG          1315
Asp Ser Tyr Tyr Glu Arg Glu Ala Asp Gly Gly Tyr Ser Asn Leu Met
            345                 350                 355
```

FIG. 14(E)

```
TTC GCC AAC GCC GCC GTG AAC TGC CTC GAC CTC CCC GCC GCC TTC TCC          1363
Phe Ala Asn Ala Ala Val Asn Cys Leu Asp Leu Pro Ala Ala Phe Ser
        360                 365                 370

TCC CCG GAC GAG GTG CGC GAC GCC CTC CCC GAC TTC GAG AAG GCG TCC          1411
Ser Pro Asp Glu Val Arg Asp Ala Leu Pro Asp Phe Glu Lys Ala Ser
    375                 380                 385

CCG GTC TTC GGC GAG GGC CTC GCC TGG TCC TCC CTG AAC TGC GCG TAC          1459
Pro Val Phe Gly Glu Gly Leu Ala Trp Ser Ser Leu Asn Cys Ala Tyr
390                 395                 400                 405

TGG CCG GTG AAG CCC ACG GGG GAG CCG CAC CGC ATC GAG GCG GCC GGC          1507
Trp Pro Val Lys Pro Thr Gly Glu Pro His Arg Ile Glu Ala Ala Gly
                410                 415                 420

GCC ACC CCG ATC GTC GTG GTC GGC ACC ACC CGC GAC CCG GCC ACC CCC          1555
Ala Thr Pro Ile Val Val Val Gly Thr Thr Arg Asp Pro Ala Thr Pro
            425                 430                 435
```

FIG. 14(F)

```
TAC CGC TGG GCC GAG GCC CTC TCC GAC CAG CTC ACC TCC GGC CAC CTC          1603
Tyr Arg Trp Ala Glu Ala Leu Ser Asp Gln Leu Thr Ser Gly His Leu
        440                 445                 450

CTC ACC TAC GAG GGA GAC GGC CAC ACC GCG TAC GGC CGC GGC AGC TCC          1651
Leu Thr Tyr Glu Gly Asp Gly His Thr Ala Tyr Gly Arg Gly Ser Ser
    455                 460                 465

TGC ATC GAC TCC GCG ATC AAC ACG TAC CTG CTG ACC GGC ACC GCC CCG          1699
Cys Ile Asp Ser Ala Ile Asn Thr Tyr Leu Leu Thr Gly Thr Ala Pro
470                 475                 480                 485

GAG GAC GGC AAG CGC TGC TCG TAA CCCCC GCCTGCCCGC CCCGGGACCC ACGCCTCCGG  1758
Glu Asp Gly Lys Arg Cys Ser ter
                490

GGGCGGGTTC GGAGCACCCC GGGAAACTGT GTAGACTTGC CGACGTTGCT GATCGCACCA TGG   1821
```

FIG. 15(A)

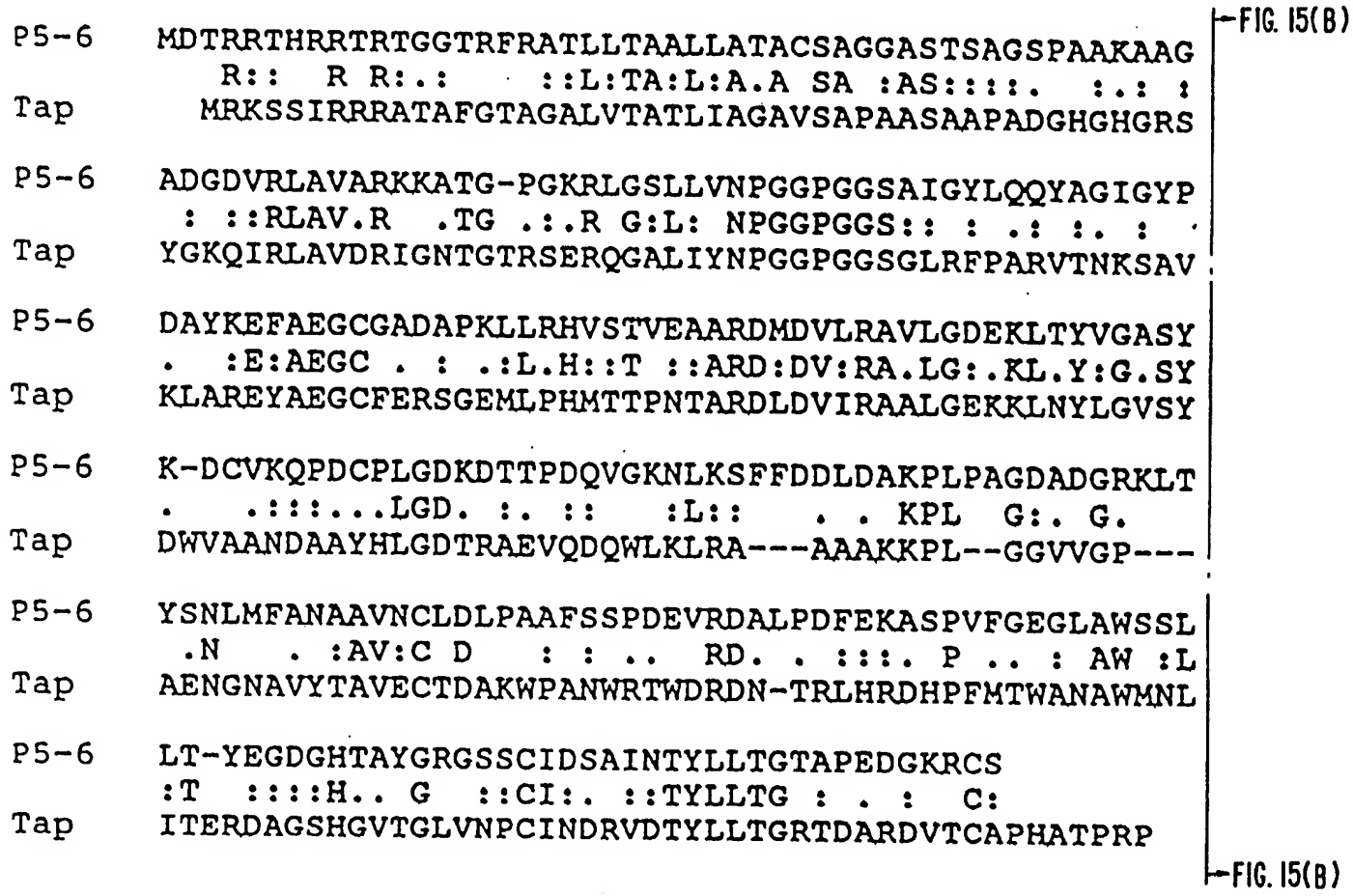


FIG. 15(B)

FIG. 15(B)

FIG. 15(B)

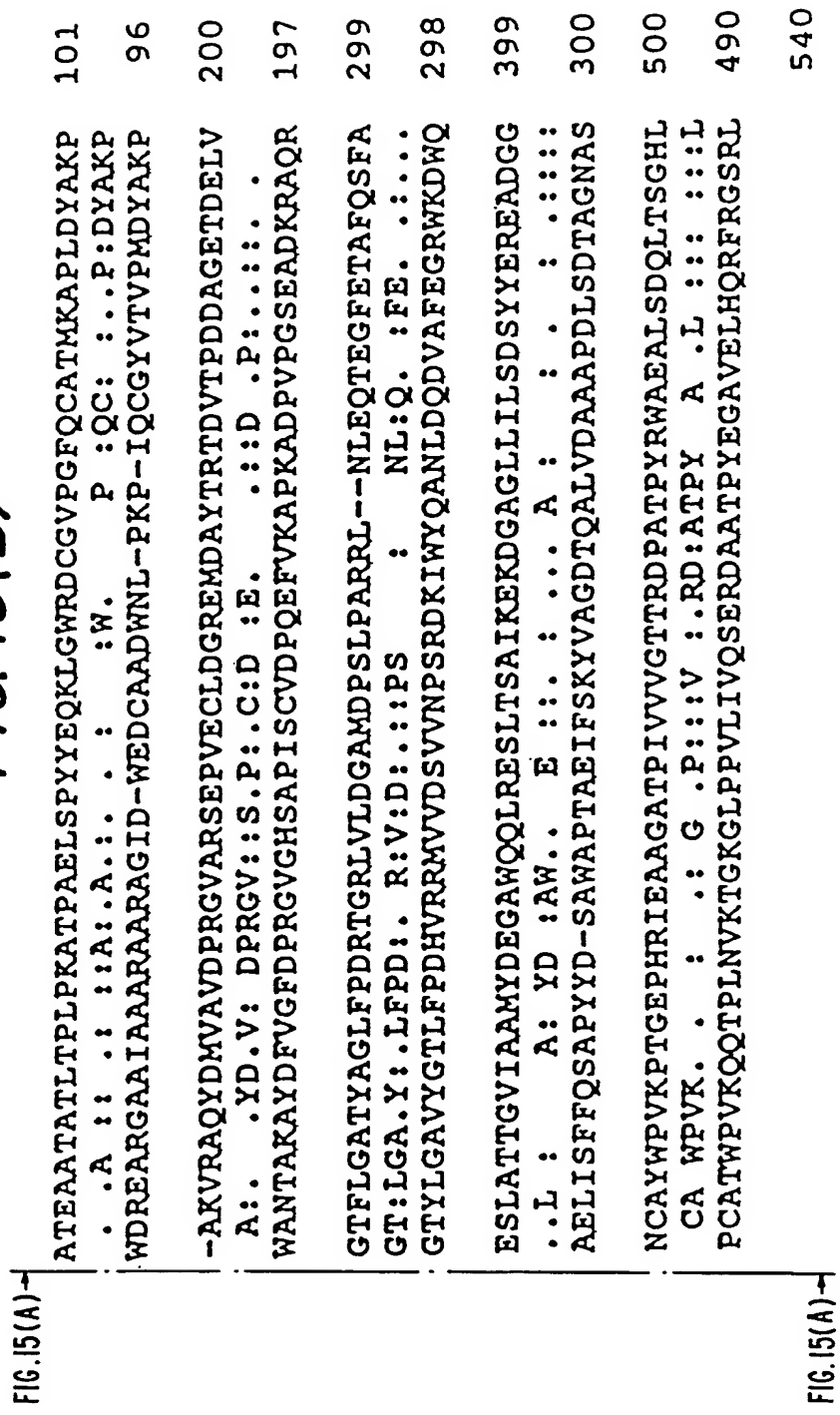


FIG.15(A)
ATEAATATLTPLPKATPAELSPYYEQKLGWRDCGVPGFQCATMKAPLDYAKP 101
. .A :: .: ::A:.A.:. . : :W. P :QC: :..P:DYAKP
WDREARGAAIAAARAARAGID-WEDCAADWNL-PKP-IQCGYVTVPMDYAKP 96
-AKVRAQYDMVAVDPRGVARSEPVECLDGREMDAYTRTDVTPDDAGETDELV 200
A:. .YD.V: DPRGV::S.P:.C:D :E. .::D .P:..::. .
WANTAKAYDFVGFDPRGVGHSAPISCVDPQEFVKAPKADPVPGSEADKRAQR 197
GTFLGATYAGLFPDRTGRLVLDGAMDPSLPARRL--NLEQTEGFETAFQSFA 299
GT:LGA.Y:.LFPD:. R:V:D:.::PS : NL:Q. :FE. .:...
GTYLGAVYGTLFPDHVRRMVVDSVVNPSRDKIWYQANLDQDVAFEGRWKDWQ 298
ESLATTGVIAAMYDEGAWQQLRESLTSAIKEKDGAGLLILSDSYYERADGG 399
..L : A: YD :AW.. E ::. : .... A : : . : .::::
AELISFFQSAPYYD-SAWAPTAEIFSKYVAGDTQALVDAAAPDLSDTAGNAS 300
NCAYWPVKPTGEPHRIEAAGATPIVVVGTTRDPATPYRWAEALSDQLTSGHL 500
CA WPVK. . : .: G .P:::V :.RD:ATPY A .L ::: :::L
PCATWPVKQQTPLNVKTGKGLPPVLIVQSERDAATPYEGAVELHQRFRGSRL 490
540
537
FIG.15(A)

FIG. 16

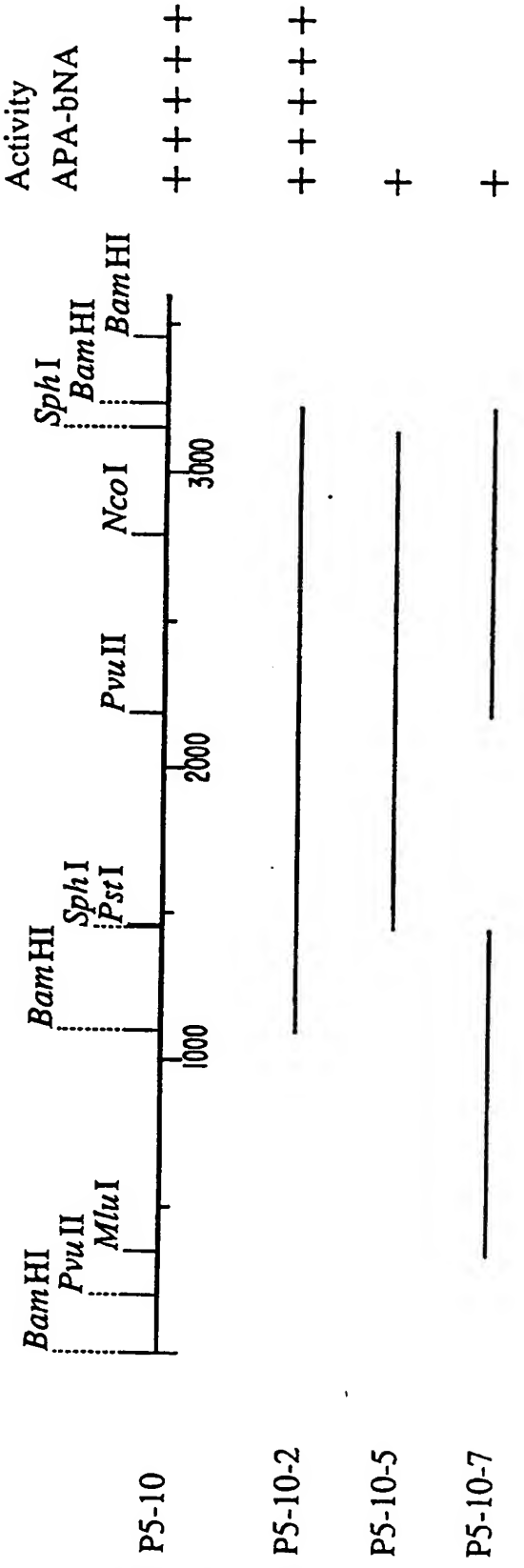

Activity
APA-bNA
+++++
+++++
+
+
BamHI
PvuII
MluI
BamHI
SphI
PstI
PvuII
NcoI
SphI
BamHI
BamHI
1000
2000
3000
P5-10
P5-10-2
P5-10-5
P5-10-7

FIG. 17(A)

```
GGATCCTCTG CTGACTCTGG TGTCGCGGGA GTTGCTCCGA TAGCTTCCXG CGGACATCAG        60

TTTCACAGGG AGGGTGC  ATG CGC AAG GCG CTC AGA TGG CTG CTG GCG CTC         110
                    fMet Arg Lys Ala Leu Arg Trp Leu Leu Ala Leu
                                      5                  10

GTG GTG CTC ATA GGC ACC GTC AGC ACG GCG GGG GCG GCC ACC GCC GCC         158
Val Val Leu Ile Gly Thr Val Ser Thr Ala Gly Ala Ala Thr Ala Ala
            15                  20                  25

GAG CCG AAG GCC GTC GAC ATC AAG GAC CGG CTG CTG TCC ATA CCG GGC         206
Glu Pro Lys Ala Val Asp Ile Lys Asp Arg Leu Leu Ser Ile Pro Gly
        30                  35                  40

ATG AGC CTG ATC GAG GAG AAG CCG TAC ACC GGC TAC CGC TTC TTC GTC         254
Met Ser Leu Ile Glu Glu Lys Pro Tyr Thr Gly Tyr Arg Phe Phe Val
    45                  50                  55

CTC AAC TAC ACC CAG CCG GTG GAC CAC CGG CAC CCG TCC AAG GGC ACG         302
Leu Asn Tyr Thr Gln Pro Val Asp His Arg His Pro Ser Lys Gly Thr
60                  65                  70                  75

TTC CAG CAG CGG ATC ACC GTG CTG CAC AAG GAC GTG AAC CGC CCG ACG         350
Phe Gln Gln Arg Ile Thr Val Leu His Lys Asp Val Asn Arg Pro Thr
                80                  85                  90

GTC TTC TAC ACC GGC GGC TAC AAC GTC TCC ACG AAC CCC AGC CGG CGC         398
Val Phe Tyr Thr Gly Gly Tyr Asn Val Ser Thr Asn Pro Ser Arg Arg
            95                  100                 105
```

FIG. 17(B)

```
GAG CCG ACC CAG ATC GTG GAC GGC AAC CAG GTC TCC ATG GAG TAC CGC          446
Glu Pro Thr Gln Ile Val Asp Gly Asn Gln Val Ser Met Glu Tyr Arg
        110                 115                 120

TAC TTC ACG CCG TCC CGG CCC GCC CCG GCC GAC TGG TCC AAG CTG GAC          494
Tyr Phe Thr Pro Ser Arg Pro Ala Pro Ala Asp Trp Ser Lys Leu Asp
    125                 130                 135

ATC TGG CAG GCC GCC AGC GAC CAG CAC CGC ATC TTC AAG GCC CTC AAG          542
Ile Trp Gln Ala Ala Ser Asp Gln His Arg Ile Phe Lys Ala Leu Lys
140                 145                 150                 155

CCG CTC TAC TCC AGG AAC TGG ATC TCC ACC GGC GGC TCC AAG GGC GGC          590
Pro Leu Tyr Ser Arg Asn Trp Ile Ser Thr Gly Gly Ser Lys Gly Gly
                160                 165                 170

ATG ACC GCC ACC TAC TAC GAG CGC TTC TAT CCG CGT GAC ATG GAC GGC          638
Met Thr Ala Thr Tyr Tyr Glu Arg Phe Tyr Pro Arg Asp Met Asp Gly
            175                 180                 185

GTC GTC GCC TAC GTC GCC CCC AAC GAC GTG GTG AAC AAG GAG GAC TCG          686
Val Val Ala Tyr Val Ala Pro Asn Asp Val Val Asn Lys Glu Asp Ser
        190                 195                 200

GCC TAC GAC CGC TTC TTC GCC CGC GTC GGC ACC GAC GAG TGC CGC GAC          734
Ala Tyr Asp Arg Phe Phe Ala Arg Val Gly Thr Asp Glu Cys Arg Asp
    205                 210                 215
```

FIG. 17(C)

```
AAG CTC AAC GGC GTG CAG CGC GAG GCG CTG GTG CGC CGG GCG CCG CTG          782
Lys Leu Asn Gly Val Gln Arg Glu Ala Leu Val Arg Arg Ala Pro Leu
220                 225                 230                 235

GAG AAG AAG TAC GCG GCG TAC GCG GCC GAG AAC GGC TAC ACC TTC GAC          830
Glu Lys Lys Tyr Ala Ala Tyr Ala Ala Glu Asn Gly Tyr Thr Phe Asp
                240                 245                 250

ACC ATC GGC AGC CTC GAC CGC GCC TAC GAG GCC GTC GTC CTC GAC TAC          878
Thr Ile Gly Ser Leu Asp Arg Ala Tyr Glu Ala Val Val Leu Asp Tyr
            255                 260                 265

GTG TGG GGC TTC TGG CAG TAC AGC ACC CTC GCC GAC TGC GCC GAC ATC          926
Val Trp Gly Phe Trp Gln Tyr Ser Thr Leu Ala Asp Cys Ala Asp Ile
        270                 275                 280

CCG GCC GAC GCC AAG AAC GCC ACC GAC GAC GCG ATC TGG GGC TCC GTC          974
Pro Ala Asp Ala Lys Asn Ala Thr Asp Asp Ala Ile Trp Gly Ser Val
    285                 290                 295

GAC GCG ATC TCC GGC TTC TCC GCC TAC ACG GAC CAG GGC CTG GAG ACG         1022
Asp Ala Ile Ser Gly Phe Ser Ala Tyr Thr Asp Gln Gly Leu Glu Thr
300                 305                 310                 315

TAC ACG CCG TAC TAC TAC CAG GCG GGC ACC CAG CTG GGC GCG CCG ACG         1070
Tyr Thr Pro Tyr Tyr Tyr Gln Ala Gly Thr Gln Leu Gly Ala Pro Thr
                320                 325                 330
```

FIG. 17(D)

```
ATC CAC TTC CCG CAC ATC GAG AAG AAG TAC ATC CGC TAC GGC TAC CAG          1118
Ile His Phe Pro His Ile Glu Lys Lys Tyr Ile Arg Tyr Gly Tyr Gln
                335                 340                 345

CCG CCG CGC AAC TTC GTG CCC CGC TCG ATC CCG ATG AAG TTC GAG CCG          1166
Pro Pro Arg Asn Phe Val Pro Arg Ser Ile Pro Met Lys Phe Glu Pro
        350                 355                 360

TGG GCG ATG CGG GAC GTC GAC ACC TGG GTG CGG CAC AAC GCC CGG CAC          1214
Trp Ala Met Arg Asp Val Asp Thr Trp Val Arg His Asn Ala Arg His
    365                 370                 380

ATG CTG TTC GTG TAC GGC GAG AAC GAC CCG TGG GGC GCC GAG CGC TTC          1262
Met Leu Phe Val Tyr Gly Glu Asn Asp Pro Trp Gly Ala Glu Arg Phe
385                 390                 395                 400

CGC CTC GGC CAC GGC GCG CGT GAC TCC TAC GTC CTC ACC GCG CCC GGC          1310
Arg Leu Gly His Gly Ala Arg Asp Ser Tyr Val Leu Thr Ala Pro Gly
                405                 410                 415

ATG AAC CAC GGT GCG AAC GTG GCC GGT CTG GTG CCC GAC CAG AAG GCG          1358
Met Asn His Gly Ala Asn Val Ala Gly Leu Val Pro Asp Gln Lys Ala
            420                 425                 430

CGG GCC ACG GGG CGC ATC CTG GAC TGG GCG GGC GTC GCC CCG GCC AAG          1406
Arg Ala Thr Gly Arg Ile Leu Asp Trp Ala Gly Val Ala Pro Ala Lys
        435                 440                 445

GTC CAG GAG AAC CCG TCG GCG GCC AGG CCG CTG GCG ACG TTC GAC GCC          1454
Val Gln Glu Asn Pro Ser Ala Ala Arg Pro Leu Ala Thr Phe Asp Ala
    450                 455                 460
```

FIG. 17(E)

```
AGG CTG GAC CAG CGG GAC GTC GAG CGC GAG CCG GCG CTG CGC CCG TAG      1502
Arg Leu Asp Gln Arg Asp Val Glu Arg Glu Pro Ala Leu Arg Pro Stop
465                 470                 475

GGCCTGCTCC TCCGGCACGG TGACTGTGTC CGGCCGCTCA CAGCGGCCGG GCACATCCCA   1562
CCGGGCGCTC GCCGCCGAGC CGGACGTACA GACCGGTCGA GGTCGGGCAC TGAGCCCGGG   1622
TGGCGACCGC CTCGACCACC TTGAACTGCG GTGCGTACCC GCCCCCGCCG GCGCAGGCCG   1682
CCTCGCGGAC CTTGCCGCCG CCGATGTCGT GGACGCAGTC GCCGACGACG GTGCGCGGAC   1742
CTCCTCCGCC GCCCGGATCA CCGGGATGCG GCGGCTGCAG CTTGCGCATG C            1793
```

FIG. 18

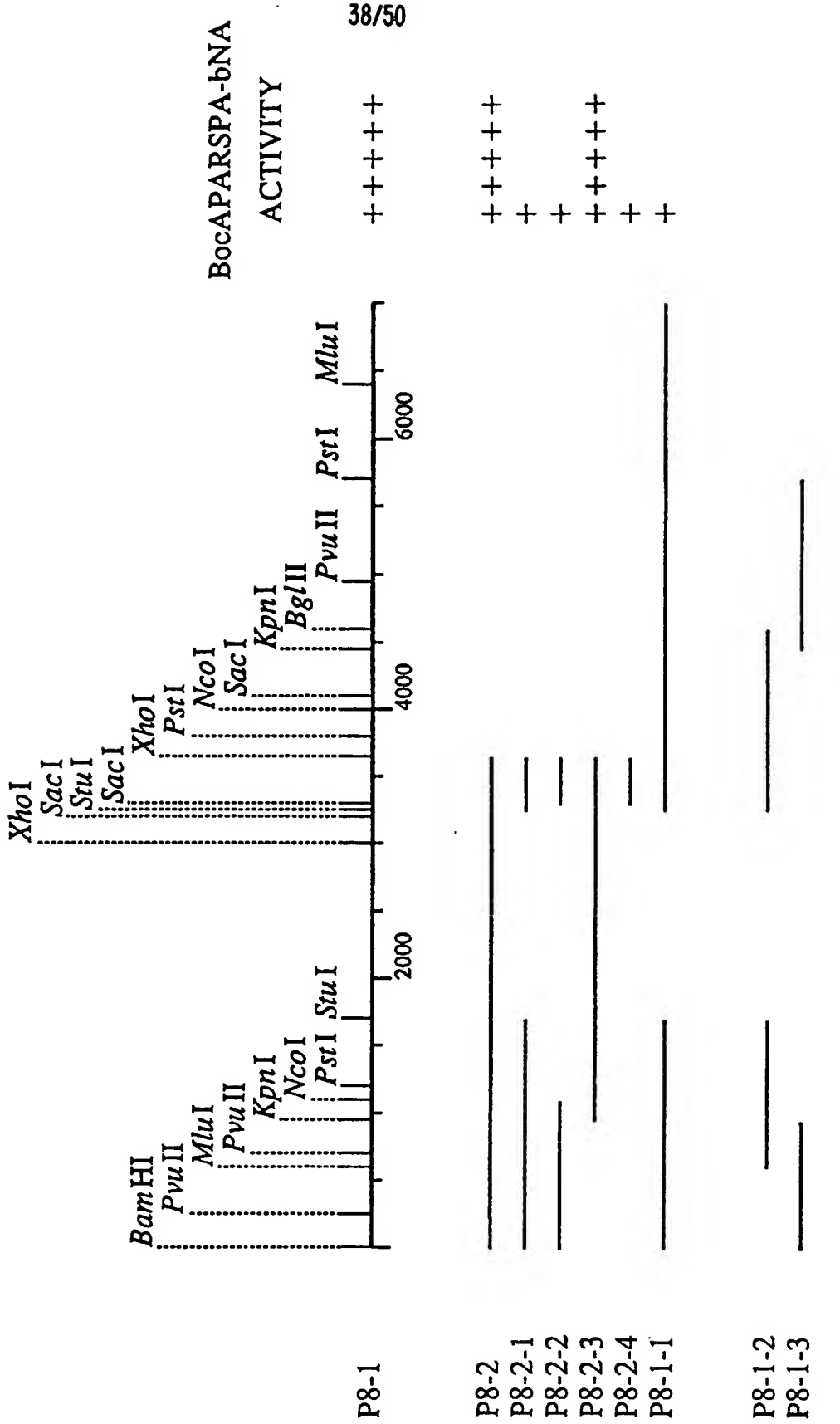

BocAPARSPA-bNA
ACTIVITY
BamHI
PvuII
MluI
PvuII
KpnI
NcoI
PstI
StuI
XhoI
SacI
StuI
SacI
XhoI
PstI
NcoI
SacI
KpnI
BglII
PvuII
PstI
MluI
2000
4000
6000
P8-1
P8-2
P8-2-1
P8-2-2
P8-2-3
P8-2-4
P8-1-1
P8-1-2
P8-1-3

FIG. 19(A)

```
GGTACCGGCG GCCAAGACCG TGTGCTCCTG ACCGCGGACG CCACCACAGG TCGGCAGAAG        60
CAGCAGATCG ACAGAAGTAG CAGGTCAGAG CGTTATCCAC AGGCGTCGGC GGGTGCTGCC       120
CCCGCCACCT ACCATGGCAG GAACGCCATC CGCCGCACGG CGCGGACGGC TTGCCAGGGG       180
GGAGAGGAC ATG GCG CGT CTC GTC CGG TGG ACG GCT CTG ACG GCC GCC GCC GCA    234
          fMet Ala Arg Leu Val Arg Trp Thr Ala Leu Thr Ala Ala Ala Ala
                                  5                  10                  15

CTG CTG ACG GCG GGC TGC AGC GGC GGC TCG TCC GAC GAG GAC AAG GAC         282
Leu Leu Thr Ala Gly Cys Ser Gly Gly Ser Ser Asp Glu Asp Lys Asp
                    20                  25                  30

GAC GGG GGC AGG AGC AGC GCG GGA CCT TCG GCG GCG GCA CCC TCC GGG         330
Asp Gly Gly Arg Ser Ser Ala Gly Pro Ser Ala Ala Ala Pro Ser Gly
                35                  40                  45

GTG CCG GAG GCA CTG GCG TCC CAG ACG CTG GAC TGG GCC CGA TGC GAG         378
Val Pro Glu Ala Leu Ala Ser Gln Thr Leu Asp Trp Ala Arg Cys Glu
            50                  55                  60

GGC AGC GAC GAT GCC CCG GCG CCG GAC GGC GAC TGG CGG TGC GCC ACG         426
Gly Ser Asp Asp Ala Pro Ala Pro Asp Gly Asp Trp Arg Cys Ala Thr
        65                  70                  75
```

FIG. 19(B)

```
CTG AAG GCA CCG CTG GAC TGG TCC GAC CCC GAC GGC GAG ACG ATC GAT          474
Leu Lys Ala Pro Leu Asp Trp Ser Asp Pro Asp Gly Glu Thr Ile Asp
 80                  85                  90                  95

CTC GCG CTG ATC CGG TCC CGG GCG AGC GGG GAC GAC CGC ATC GGC TCC          522
Leu Ala Leu Ile Arg Ser Arg Ala Ser Gly Asp Asp Arg Ile Gly Ser
                100                 105                 110

CTG CTG TTC AAC TTC GGC GGC CCG GGC GCC TCC GGC GTC TCC ACG ATG          570
Leu Leu Phe Asn Phe Gly Gly Pro Gly Ala Ser Gly Val Ser Thr Met
            115                 120                 125

CCG TCC TAC GCC GAC ACC GTC TCC TCC CTG CAC GAG CGG TAC GAC CTG          618
Pro Ser Tyr Ala Asp Thr Val Ser Ser Leu His Glu Arg Tyr Asp Leu
        130                 135                 140

GTG AGC TGG GAC CCG CGC GGG GTG GCC GCC AGC GAG GGC GTC CGC TGC          666
Val Ser Trp Asp Pro Arg Gly Val Ala Ala Ser Glu Gly Val Arg Cys
    145                 150                 155

CGC ACC GAC GAG GCG ATC GAG GCC GCC GAG TCG GTG GAC TCC ACG CCG          714
Arg Thr Asp Glu Ala Ile Glu Ala Ala Glu Ser Val Asp Ser Thr Pro
160                 165                 170                 175

GAC TCC CCG GCC GAG GAG CAG GCC TAC CTG AAG GAC GCC GCC GAC TTC          762
Asp Ser Pro Ala Glu Glu Gln Ala Tyr Leu Lys Asp Ala Ala Asp Phe
                180                 185                 190
```

FIG. 19(C)

```
GGC AGG GGC TGC GAG AAG GCC GCC GGC AAG CTC ATG GAA CAC GTC TCG          810
Gly Arg Gly Cys Glu Lys Ala Ala Gly Lys Leu Met Glu His Val Ser
            195                 200                 205

ACC ACG GAC ACG GCC CGC GAC ATG GAC CTG ATG CGG CAC GTC CTG GGC          858
Thr Thr Asp Thr Ala Arg Asp Met Asp Leu Met Arg His Val Leu Gly
        210                 215                 220

GAC GAG AGG ATG CAC TAC TTC GGC ATC TCC TAC GGC ACC GAA CTC GGC          906
Asp Glu Arg Met His Tyr Phe Gly Ile Ser Tyr Gly Thr Glu Leu Gly
    225                 230                 235

GGC GTC TAC GCC CAT CTG TTC CCC GAG CAC GTG GGC CGC GTG ATC CTC          954
Gly Val Tyr Ala His Leu Phe Pro Glu His Val Gly Arg Val Ile Leu
240                 245                 250                 255

GAC GCG GTG GTG GAC CCG GGC GCC GAC ACG ATG GGC CAC GCC GAG AAC         1002
Asp Ala Val Val Asp Pro Gly Ala Asp Thr Met Gly His Ala Glu Asn
                260                 265                 270

CAG GCC AGG GGT TTC CAG CGC GCG CTG GAC GAC TAC CTG GAG TCG ACC         1050
Gln Ala Arg Gly Phe Gln Arg Ala Leu Asp Asp Tyr Leu Glu Ser Thr
            275                 280                 285

GGC CAG GAA CCC GAA CAG GGG TCG CGG AAG ATC GCC GGC CTG CTG GAG         1098
Gly Gln Glu Pro Glu Gln Gly Ser Arg Lys Ile Ala Gly Leu Leu Glu
        290                 295                 300
```

FIG. 19(D)

```
CGG CTG GAC GCC GAG CCA CTG CCC ACG TCC TCG CCG GGG CGG GAG CTG        1146
Arg Leu Asp Ala Glu Pro Leu Pro Thr Ser Ser Pro Gly Arg Glu Leu
    305                 310                 315

ACG CAG ACC CTC GCG TTC ACC GGC ATC GTG CTG CCG CTG TAC AGC GAG        1194
Thr Gln Thr Leu Ala Phe Thr Gly Ile Val Leu Pro Leu Tyr Ser Glu
320                 325                 330                 335

AGC GGC TGG CCG GCC CTG ACC AGT GCG CTG AAG GCG GCC GAG GAG GGC        1242
Ser Gly Trp Pro Ala Leu Thr Ser Ala Leu Lys Ala Ala Glu Glu Gly
                340                 345                 350

GAC GGC TCG GAG TTG CTG GCC CTC GCC GAC GGC TAC AAC GAG CGT GAT        1290
Asp Gly Ser Glu Leu Leu Ala Leu Ala Asp Gly Tyr Asn Glu Arg Asp
            355                 360                 365

CCC TCG GGG CGC TAC GGC ACG ACG ACC CAC TCG CAA AGG GTC ATA TCG        1338
Pro Ser Gly Arg Tyr Gly Thr Thr Thr His Ser Gln Arg Val Ile Ser
        370                 375                 380

TCG CTG GAC GAC AAG CAG AGG CCG ACC GTG GAG GAG ACG AAG AAG CTG        1386
Cys Leu Asp Asp Lys Gln Arg Pro Thr Val Glu Glu Thr Lys Lys Leu
    385                 390                 395
```

FIG. 19(E)

```
CTG CCG AGG TTC GAG AAG GTC TCT CCC GTC TTC GGC GCC TTC CTC GGC      1434
Leu Pro Arg Phe Glu Lys Val Ser Pro Val Phe Gly Ala Phe Leu Gly
400                 405                 410                 415

TGG GAC ACG GCC GGG TGG TGC CAC GAC TGG CCG GTG GCC GGT CAG CAC      1482
Trp Asp Thr Ala Gly Trp Cys His Asp Trp Pro Val Ala Gly Gln His
                420                 425                 430

GAG ACC GCG GAG GTG AGC GCG CCC GAC GCG GCC CCG GTC CTG GTG GTC      1530
Glu Thr Ala Glu Val Ser Ala Pro Asp Ala Ala Pro Val Leu Val Val
            435                 440                 445

GGC AAC ACG GGC GAC CCG GCC ACG CCC TAC GAG GGC GCC CGC AGG ATG      1578
Gly Asn Thr Gly Asp Pro Ala Thr Pro Tyr Glu Gly Ala Arg Arg Met
        450                 455                 460

GCG GAC GAG CTG GGC AAG GAC GTC GGC GTG GTG CTG ACC TGG CAG GGC      1626
Ala Asp Glu Leu Gly Lys Asp Val Gly Val Val Leu Thr Trp Gln Gly
    465                 470                 475  .

GAG GGA CAC GGT GCC TAC GGG AAC GGA AGC GAC TGT GTC GAC TCC GCG      1674
Glu Gly His Gly Ala Tyr Gly Asn Gly Ser Asp Cys Val Asp Ser Ala
480                 485                 490                 495

GTG GAC GCC TAC CTG TTG AAG GGG ACG GTG CCG AAG GAC GGC AAG GTC      1722
Val Asp Ala Tyr Leu Leu Lys Gly Thr Val Pro Lys Asp Gly Lys Val
        500                 505                 510

TGC TCA TGA CGGCGGCGGG GGCTTCGGGC ACCTGCGGTG CGCGAAACCC CCGCCG       1771
Cys Ser End
```

FIG. 20(B)

```
GGTCAACATCGTGGACGACCTGGTGGAGTGCGTCAAGGAGAACTCGAGCAAGGACCTCAA
----.----+----.----+----.----+----.----+----.----+----.----+ 300
CCAGTTGTAGCACCTGCTGGACCACCTCACGCAGTTCCTCTTGAGCTCGTTCCTGGAGTT

GAAGAGCTTCAAGTCCCCCGAGCCCCGCCTGTTCACGCCCGAGGAGTTCTTCCGGATCTT
----.----+----.----+----.----+----.----+----.----+----.----+ 360
CTTCTCGAAGTTCAGGGGGCTCGGGGCGGACAAGTGCGGGCTCCTCAAGAAGGCCTAGAA

CAACCGCTCGATCGACGCCTTCAAGGACTTCGTCGTGGCGTCCGAGACCTCCGACTGCGT
----.----+----.----+----.----+----.----+----.----+----.----+ 420
GTTGGCGAGCTAGCTGCGGAAGTTCCTGAAGCAGCACCGCAGGCTCTGGAGGCTGACGCA

GGTCAGCTCGACCCTCAGCCCGGAGAAGGACTCGCGGGTGTCGGTCACCAAGCCGTTCAT
----.----+----.----+----.----+----.----+----.----+----.----+ 480
CCAGTCGAGCTGGGAGTCGGGCCTCTTCCTGAGCGCCCACAGCCAGTGGTTCGGCAAGTA

                    Hind III
GCTGCCCCCCGTCGCCTA
----.----+----.----+--  502
CGACGGGGGGCAGCGGATTCGA
```

FIG. 21(A)

PstI

```
   GACTGCGACATCGAGGGGAAGGACGGCAAGCAGTACGAGTCGGTGCTGATGGTGTC
----.----+----.----+----.----+----.----+----.----+----.----+   60
ACGTCTGACGCTGTAGCTCCCCTTCCTGCCGTTCGTCATGCTCAGCCACGACTACCACAG

CATCGACCAGTTGCTGGACTCGATGAAGGAGATCGGCTCCAACTGCCTCAACAACGAGTT
----.----+----.----+----.----+----.----+----.----+----.----+  120
GTAGCTGGTCAACGACCTGAGCTACTTCCTCTAGCCGAGGTTGACGGAGTTGTTGCTCAA

CAACTTCTTCAAGCGCCACATCTGCGACGCCAACAAGGAGGGAATGTTCCTGTTCCGGGC
----.----+----.----+----.----+----.----+----.----+----.----+  180
GTTGAAGAAGTTCGCGGTGTAGACGCTGCGGTTGTTCCTCCCTTACAAGGACAAGGCCCG

CGCGCGCAAGCTGCGCCAGTTCCTCAAGATGAATTCCACCGGGGACTTCGACCTCCACCT
----.----+----.----+----.----+----.----+----.----+----.----+  240
GCGCGCGTTCGACGCGGTCAAGGAGTTCTACTTAAGGTGGCCCCTGAAGCTGGAGGTGGA
```

FIG. 21(B)

```
GCTCAAGGTCTCGGAGGGCACGACCATCCTGCTGAACTGCACGGGCCAGGTCAAGGGACG
----.----+----.----+----.----+----.----+----.----+----.----+  300
CGAGTTCCAGAGCCTCCCGTGCTGGTAGGACGACTTGACGTGCCCGGTCCAGTTCCCTGC

GAAGCCCGCCGCCCTCGGGGAGGCCCAGCCGACGAAGAGCTTGGAGGAAAACAAGTCCCT
----.----+----.----+----.----+----.----+----.----+----.----+  360
CTTCGGGCGGCGGGAGCCCCTCCGGGTCGGCTGCTTCTCGAACCTCCTTTTGTTCAGGGA

GAAGGAGCAGAAGAAGCTCAACGACCTGTGCTTCCTGAAGCGGTTGCTCCAGGAGATCAA
----.----+----.----+----.----+----.----+----.----+----.----+  420
CTTCCTCGTCTTCTTCGAGTTGCTGGACACGAAGGACTTCGCCAACGAGGTCCTCTAGTT

                                                  HindIII
GACGTGCTGGAACAAGATCCTCATGGGCACCAAGGAACACTGA
----.----+----.----+----.----+----.----+----.--   467
CTGCACGACCTTGTTCTAGGAGTACCCGTGGTTCCTTGTGACTTCGA
```

FIG. 22(A)

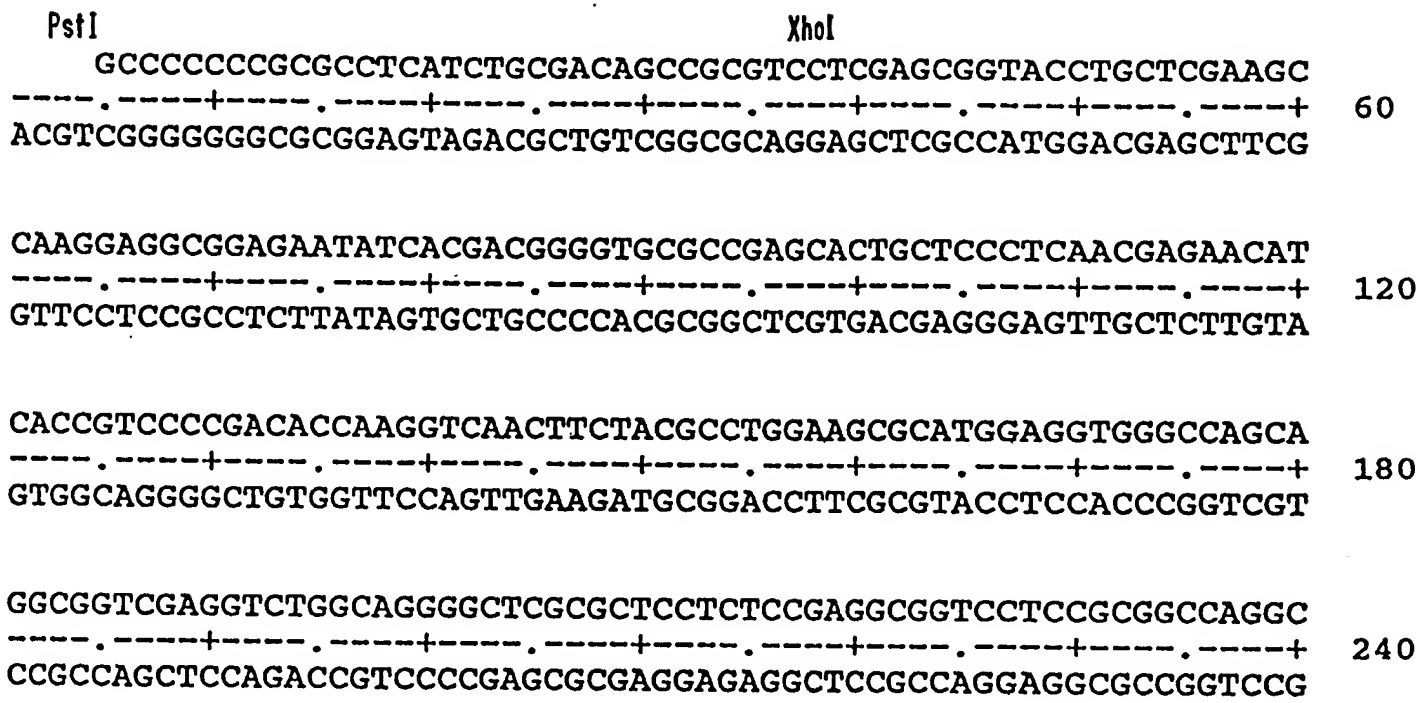

PstI
XhoI
GCCCCCCCGCGCCTCATCTGCGACAGCCGCGTCCTCGAGCGGTACCTGCTCGAAGC
60
ACGTCGGGGGGGCGCGGAGTAGACGCTGTCGGCGCAGGAGCTCGCCATGGACGAGCTTCG
CAAGGAGGCGGAGAATATCACGACGGGGTGCGCCGAGCACTGCTCCCTCAACGAGAACAT
120
GTTCCTCCGCCTCTTATAGTGCTGCCCCACGCGGCTCGTGACGAGGGAGTTGCTCTTGTA
CACCGTCCCCGACACCAAGGTCAACTTCTACGCCTGGAAGCGCATGGAGGTGGGCCAGCA
180
GTGGCAGGGGCTGTGGTTCCAGTTGAAGATGCGGACCTTCGCGTACCTCCACCCGGTCGT
GGCGGTCGAGGTCTGGCAGGGGCTCGCGCTCCTCTCCGAGGCGGTCCTCCGCGGCCAGGC
240
CCGCCAGCTCCAGACCGTCCCCGAGCGCGAGGAGAGGCTCCGCCAGGAGGCGCCGGTCCG

FIG. 22(B)

```
CCTCCTGGTGAACTCGTCCCAGCCGTGGGAGCCGCTCCAGCTGCACGTCGACAAGGCCGT
----.----+----.----+----.----+----.----+----.----+----.----+  300
GGAGGACCACTTGAGCAGGGTCGGCACCCTCGGCGAGGTCGACGTGCAGCTGTTCCGGCA

CTCCGGGCTCCGGTCCCTGACCACGCTGCTGCGCGCCCTCGGTGCCCAGAAGGAGGCCAT
----.----+----.----+----.----+----.----+----.----+----.----+  360
GAGGCCCGAGGCCAGGGACTGGTGCGACGACGCGCGGGAGCCACGGGTCTTCCTCCGGTA

CTCGCCCCCGGACGCCGCCAGCGCCGCCCCGCTGCGGACGATCACGGCGGACACCTTCCG
----.----+----.----+----.----+----.----+----.----+----.----+  420
GAGCGGGGGCCTGCGGCGGTCGCGGCGGGGCGACGCCTGCTAGTGCCGCCTGTGGAAGGC

CAAGCTGTTCCGGGTCTACTCGAACTTCCTGCGGGGGAAGCTGAAGCTCTACACCGGCGA
----.----+----.----+----.----+----.----+----.----+----.----+  480
GTTCGACAAGGCCCAGATGAGCTTGAAGGACGCCCCCTTCGACTTCGAGATGTGGCCGCT

StuI                          HindIII
GGCCTGCCGCACGGGCGACCGGTA
----.----+----.----+----.---  508
CCGGACGGCGTGCCCGCTGGCCATTCGA
```

FIG. 23(A)

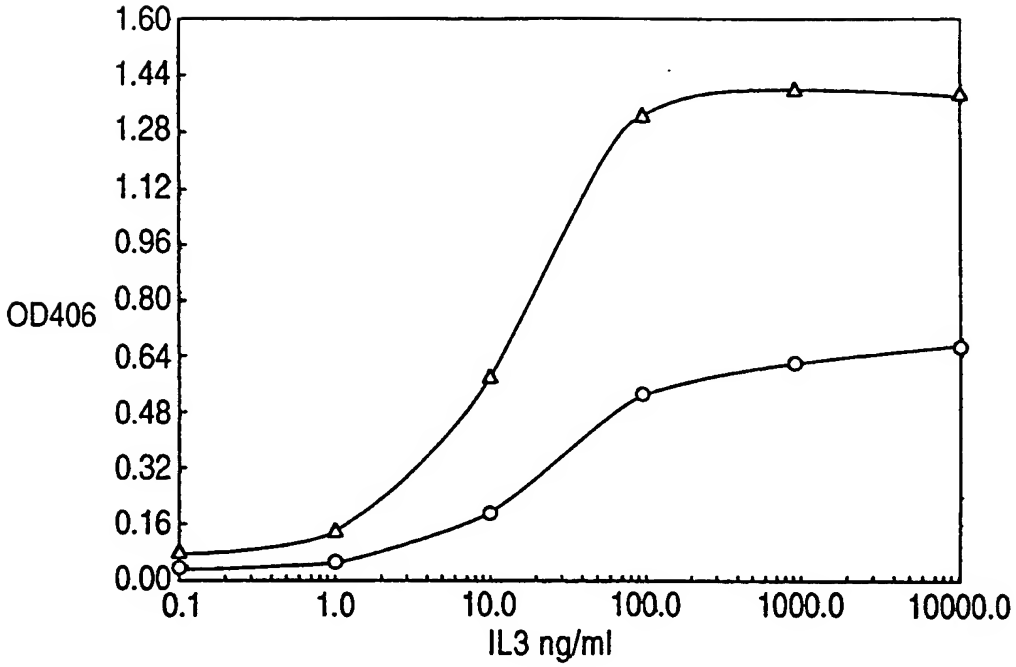


FIG. 23(B)

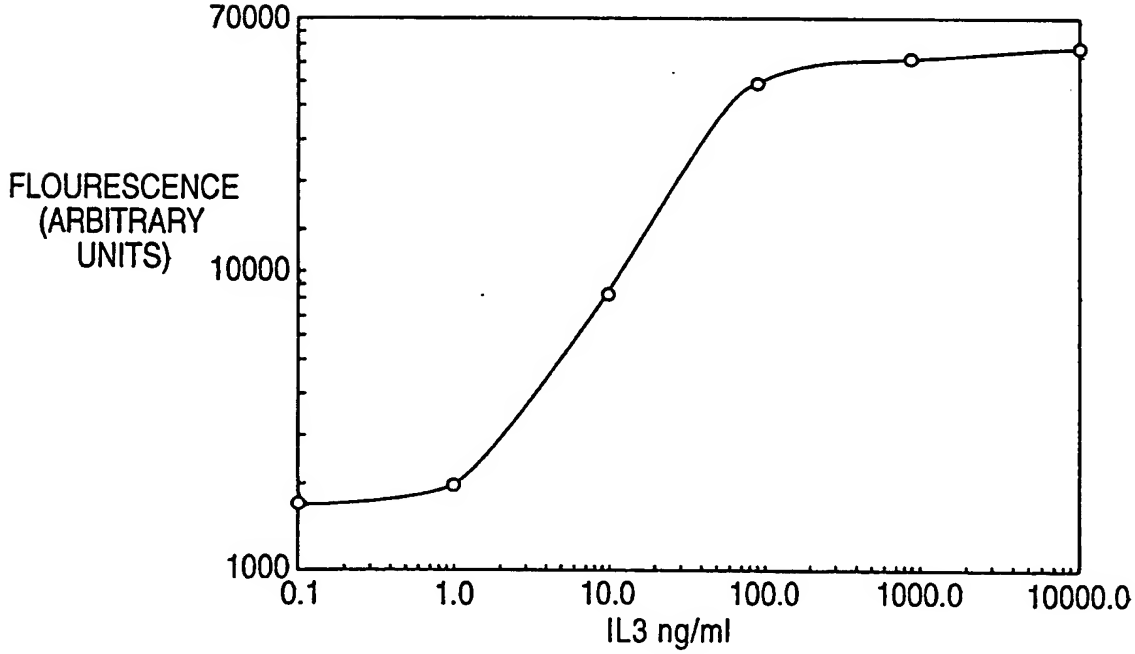
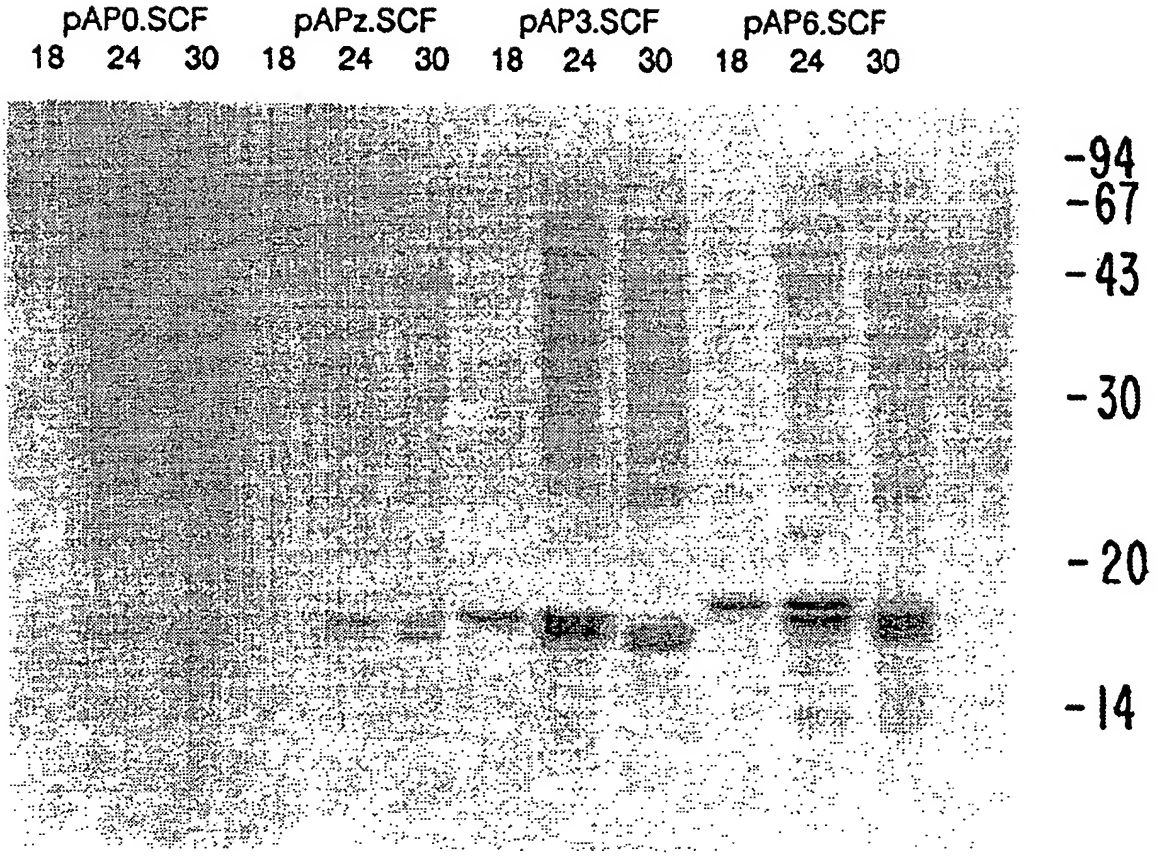


FIG. 24


pAP0.SCF
pAPz.SCF
pAP3.SCF
pAP6.SCF
18 24 30 18 24 30 18 24 30 18 24 30
-94
-67
-43
-30
-20
-14

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/57, 9/52, 9/48, 15/76, 1/21, 1/15, C07K 5/06, C12N 15/67, 15/62, C07K 14/47, C12P 21/06, G01N 33/535 // (C12N 1/21, C12R 1:465)

A3

(11) International Publication Number: **WO 95/17512**

(43) International Publication Date: 29 June 1995 (29.06.95)

(21) International Application Number: PCT/US94/14772

(22) International Filing Date: 22 December 1994 (22.12.94)

(30) Priority Data:
08/173,508 23 December 1993 (23.12.93) US

(60) Parent Application or Grant
(63) Related by Continuation
US 08/173,508 (CIP)
Filed on 23 December 1993 (23.12.93)

(71) Applicant *(for all designated States except US)*: CANGENE CORPORATION [CA/CA]; 6280 Northwest Drive, Mississauga, Ontario L4V 1J7 (CA).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: BARTFELD, Daniel [IL/CA]; 89 Overlook Place, North York, Ontario M3H 4P5 (CA). BUTLER, Michael, J. [GB/GB]; Worts Causeway, Cambridge CB1 4RN (GB). HADARY, Dany [IL/CA]; 40 Maryvale Crescent, Richmond Hill, Ontario L4C 6P8 (CA). JENISH, David, L. [CA/CA]; 4104 Fieldgate Drive, Mississauga, Ontario L4W 2C4 (CA). KRIEGER, Timothy, J. [US/CA]; 24 Jameson Crescent, Brampton, Ontario L6S 3W3 (CA). MALEK, Lawrence, T. [US/CA]; 3 Viewmount Crescent, Brampton, Ontario L6Z 4P4 (CA). WALCYZK, Eva [CA/CA]; 6037 Childham Crescent, Mississauga, Ontario L5N 2R8 (CA). SOOSTMEYER, Gisela [CA/CA]; 166 Hedgerow Lane, Kleinburg, Ontario L0G 1C0 (CA).

(74) Agents: BENT, Stephen, A. et al.; Foley & Lardner, Suite 500, 3000 K Street, N.W., Washington, DC 20007-5109 (US).

(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LT, LU, LV, MD, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report:
28 December 1995 (28.12.95)

(54) Title: PROTEASES FROM STREPTOMYCES AND USE THEREOF IN PROTEIN EXPRESSION SYSTEMS

(57) Abstract

A family of proteases endogenous to *Streptomyces* cells degrades exogenous proteins secreted from *Streptomyces* host cells. The previously unidentified proteases include (1) tripeptidyl aminopeptidase designated "Tap", (2) tripeptidyl aminopeptidase designated "Ssp", (3) X-Pro-Metalloendoproteinase designated "XP-Mep", and (4) other proteases derived from *Streptomyces* which degrade certain substrates under certain conditions. Degradation was alleviated by selective inhibition of secreted proteases or by using improved strains which lack or have impaired degradation proteases. An irreversible inhibitor was designed based upon the mechanism and substrate specificity of the target protease. Hosts expressing proteases were also produced. Uses of the proteases include immunoassays and proteolytic removal of peptides and polypeptides to improve secretion of exogenous proteins.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

INTERNATIONAL SEARCH REPORT

Int. onal Application No
PCT/US 94/14772

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/57 C12N9/52 C12N9/48 C12N15/76 C12N1/21
C12N1/15 C07K5/06 C12N15/67 C12N15/62 C07K14/47
C12P21/06 G01N33/535 //(C12N1/21,C12R1:465)
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED
Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE CHEMABS CHEMICAL ABSTRACTS SERVICE, COLUMBUS, OHIO, US LOSEVA, A. L. ET AL 'Properties of immobilized complexes of Streptomyces griseus proteinases on different carriers' see abstract & UKR. BIOKHIM. ZH. (1979), 51(4), 345-9 CODEN: UBZHD4, ---	1,2
X	WO,A,93 00925 (AMGEN INC.) 21 January 1993 see the whole document --- -/--	1,6,8, 11-19, 22,24, 25,28-30

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

2

Date of the actual completion of the international search	Date of mailing of the international search report
24 July 1995	28. 11. 95

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	VAN DER SCHAAL C.A.

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Int. onal Application No
PCT/US 94/14772

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GENE, vol.111, pages 125 - 130 H. LICHENSTEIN ET AL 'Cloning and characterization of a gene encoding extracellular metalloprotease from Streptomyces lividans' see the whole document	1,6, 13-16, 18,19
X	CHEMICAL ABSTRACTS, vol. 119, no. 3, 19 July 1993, Columbus, Ohio, US; abstract no. 23336, J. APHALE AND W. STROHL 'Purification and properties of an extracellular aminopeptidase from Streptomyces lividans 1326' page 385 ; see abstract & J. GEN. MICROBIOL., vol.139, no.3, pages 417 - 424	1,6, 13-16, 18,19
A	EP,A,0 219 237 (1CETUS CORPORATION) 22 April 1987 see the whole document	32
P,X	FEBS LETTERS, vol.352, 3 October 1994 pages 385 - 388 T. KRIEGER ET AL 'Purification and characterization of a novel tripeptidyl aminopeptidase from Streptomyces lividans 66' see the whole document	1,2,4, 6-19, 22-30,40

2

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US94/14772

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

- 6 subjects See additional sheet PCT/ISA/210

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searches without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

- see additional sheet PCT/ISA/210, pt. 1

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International Application No. PCT/US94/ 14772

FURTHER INFORMATION CONTINUED FROM PCT/ISA/210

1. claims 2,4,7,9,10,20,21,23,26,27 completely and
1,6,8,11-19,22,24,25,28-40,42,43 partially

Tripeptidyl aminopeptidases from Streptomyces, impaired forms thereof, uses thereof (e.g. in expression systems or enzyme linked immunosorbent assays), DNA encoding and methods to produce them, and an inhibitor therefor.

2. claims 3,5,41 completely and
1,6,8,11-19,22,24,25,28-40,42,43 partially

Metalloendoproteinase from Streptomyces, impaired forms thereof, uses thereof (e.g. in expression systems or enzyme linked immunosorbent assays, DNA encoding and methods to produce it.

3. claims 8,11,12,17,22-30 partially

DNA encoding impaired proteases from Streptomyces other than mentioned in subject 1 or 2, expression systems containing it.

4. claim 31 partially

Kit for enzyme-linked immunosorbent assyas comprising a protease from Streptomyces other than mentioned in subject 1 or 2, covalently linked to a carrier.

5. claims 32-39 partially

Methods to improve expression and secretion of mature protein from a genetic expression system using other proteases than mentioned in subject 1-3.

6. claims 40,42,43 partially

Use of proteases from Streptomyces other than mentioned in subject 1-3 to remove an amino acid, peptide or polypeptide from a substrate.

The International Searching Authority also draws attention to the fact that any further search for claims as grouped above may rise to further objections concerning the unity of said claimed inventions within the sense of Rule 13.1 PCT.

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int. .onal Application No
PCT/US 94/14772

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
WO-A-9300925	21-01-93	AU-B-	655312	15-12-94
		AU-A-	2307592	11-02-93
		CA-A-	2090463	02-01-93
		EP-A-	0546156	16-06-93
		JP-T-	6501396	17-02-94
		NZ-A-	243356	27-04-94
EP-A-0219237	22-04-87	US-A-	4870017	26-09-89
		US-A-	4865974	12-09-89
		AU-B-	596810	17-05-90
		AU-B-	6294486	26-03-87
		JP-C-	1761398	20-05-93
		JP-B-	4048431	06-08-92
		JP-A-	62115281	26-05-87
		JP-A-	3277287	09-12-91
		US-A-	5013662	07-05-91